



# Identification passive en acoustique : estimateurs et applications au SHM

Remy Vincent

## ► To cite this version:

Remy Vincent. Identification passive en acoustique : estimateurs et applications au SHM. Traitement du signal et de l'image [eess.SP]. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAT020 . tel-01314751

**HAL Id: tel-01314751**

**<https://theses.hal.science/tel-01314751>**

Submitted on 11 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE**

Spécialité : **traitement du signal**

Arrêté ministériel : 7 août 2006

Présentée par

**Rémy VINCENT**

Thèse dirigée par **Olivier MICHEL**

et encadrée par **Mikael CARMONA**

préparée au sein du **CEA-Leti & Gipsa-lab**

dans l'école doctorale EEATS

## Identification passive en acoustique : estimateurs et applications au SHM

Thèse soutenue publiquement le **08 / 01 / 2016**,  
devant le jury composé de :

**Pierre-Olivier AMBLARD**

DR Gipsa-lab DIS/CICS UMR 5216, Président du jury

**Cédric RICHARD**

Pr. Laboratoire Lagrange UMR 7293, Rapporteur

**Pierre CHAINAIS**

MC Lagis UMR 8219, Rapporteur

**Philippe REFREGIER**

Pr. Institut Fresnel UMR 7249, Examineur

**Florent CHATELAIN**

MC Gipsa-lab DIS, Invité

**Jean-Louis LACOUME**

Pr. CEA-Leti DSIS, Invité

**François FRASSATI**

Ingénieur CEA-Leti DSIS/SCSE/LFAO, Invité

**Olivier J. J. MICHEL**

Pr. Gipsa-lab DIS/CICS, Directeur de thèse

**Mikael CARMONA**

Ingénieur-docteur CEA-Leti DSIS/SCSE/LFAO, Encadrant de thèse





UNIVERSITÉ DE GRENOBLE  
ÉCOLE DOCTORALE EEATS

T H È S E

TRAITEMENT DU SIGNAL

Rapport de thèse

Rémy VINCENT

**Identification passive en acoustique :  
estimateurs et application au SHM**

Thèse dirigée par Olivier MICHEL

préparée au CEA-Leti & Gipsa-lab

**Jury :**

<i>Président :</i>	Pierre-Olivier AMBLARD	-	DR. CNRS Gipsa-lab UMR 5216
<i>Rapporteurs :</i>	Cédric RICHARD	-	Pr. Laboratoire Lagrange UMR 7293
	Pierre CHAINAIS	-	MCf CRISTAL UMR 8219
<i>Examineur :</i>	Philippe REFREGIER	-	Pr. Institut Fresnel UMR 7249
<i>Invités :</i>	François FRASSATI	-	Ingénieur CEA-Leti DSIS
	Jean-Louis LACOUME	-	Pr. CEA-Leti / Gipsa-lab
	Florent CHATELAIN	-	MCf Gipsa-lab DIS UMR 5216
<i>Directeur :</i>	Olivier MICHEL	-	Pr. Gipsa-lab DIS/CICS UMR 5216
<i>Encadrant :</i>	Mikael CARMONA	-	Dr. CEA-Leti DSIS



# Remerciements

Je remercie sincèrement mes encadrants pour leur disponibilité et leur passion. Quand je pense au monde de la recherche, je pense à eux, et je me dis que tôt ou tard j’y reviendrai.

Je remercie l’ensemble des personnes qui ont accepté, à un moment ou à un autre de la thèse, d’évaluer mon travail. De nombreuses discussions enrichissantes en sont nées et je compte ces moments-là parmi les plus importants de mon travail.

Je remercie mes collègues et amis pour avoir construit un environnement que je regrette déjà.

Je remercie ma famille pour son soutien et ses encouragements.

Je remercie ma femme, pour son amour et sa compréhension.



# Préface — Trois ans, c'était demain.

Pendant mes études, j'ai toujours jeté un regard envieux au monde de la recherche, où l'activité naît de la capacité à se poser les bonnes questions et où la finalité est atteinte par une savante combinaison d'intuition et de savoir-faire. Pour constituer mon sujet de thèse, j'ai parcouru diverses thématiques avec un critère important en tête : je me projetais comment le domaine serait servi par le traitement du signal. Ah, quelle belle discipline que le traitement du signal ! Manier les mathématiques et la physique pour exploiter notre capacité à apprendre du monde qui nous entoure en l'observant... voilà de quoi séduire ! Le sujet que j'ai retenu regroupe des communautés scientifiques diverses et est par essence multi-disciplinaire. Il m'intriguait par ses promesses et ses challenges. La veille, Mikael et Olivier avaient en effet ouvert la porte sur une thématique passionnante où la physique était encore discutée et dont la compréhension ne cessait d'évoluer. Ma curiosité ainsi piquée, j'étais pris au piège. Nous voilà maintenant trois années plus tard, et ce qui est rapporté dans ce manuscrit est la trace de notre mouvement, un récit (heureusement non-chronologique) de notre itinéraire.

---

Le thème de recherche de ce manuscrit touche à l'identification des milieux de propagation par la modalité acoustique. L'identification, c'est-à-dire l'estimation des paramètres qui caractérisent le milieu, a lieu dans un contexte où les sources qui déclenchent une réponse de sa part ne sont pas contrôlées. Il en sera fait référence dans toute la suite comme étant le « contexte passif d'estimation ». Un des enjeux amenés par ce contexte surprenant est la capacité à exploiter les champs issus de ces sources d'opportunité, d'en extraire l'information vraisemblablement très variable qu'ils portent sur le milieu de propagation. En ce sens, notre contribution principale sur le sujet consiste en l'étude de la théorie de l'estimation/détection dans ce contexte passif. La portée applicative suivie dans nos travaux est dans la continuité de l'état de l'art de l'identification dans un contexte passif ; le Structural Health Monitoring (SHM), discipline qui s'emploie à la mesure et au suivi dans le temps de l'état de santé de bâtiments tels qu'un pont, un immeuble ou encore un avion. La modalité acoustique aux longueurs d'ondes audibles est adoptée, et se trouve être complémentaire à l'état de l'art ; elle donne accès à des paramètres (géométriques) que les autres ne voient pas. Enfin, pour aller plus loin vers l'application et développer des algorithmes cohérents avec un contexte applicatif réaliste, un noeud de mesure qui aurait cette fonction d'évaluateur d'état de santé d'une structure est décrit à un niveau fonctionnel, générique mais complet, tenant compte des critères d'autonomie intrinsèques au calcul embarqué sur un réseau de capteurs.

Ce manuscrit articule nos contributions sur ces trois aspects, en cours d'évaluation ou déjà acceptées par la communauté scientifique.





# Table des matières

<b>1</b>	<b>Introduction à l'identification passive</b>	<b>1</b>
1.1	Exemple introductif : l'oscillateur amorti . . . . .	3
1.2	Historique, état de l'art et positionnements . . . . .	8
1.3	Travaux présentés . . . . .	19
<b>2</b>	<b>Modèles, estimation et détection pour l'identification passive</b>	<b>21</b>
2.1	Identification passive en acoustique aérienne . . . . .	23
2.2	Théorie de l'estimation/détection pour le contexte passif . . . . .	27
2.3	Expérimentation en acoustique confinée . . . . .	32
2.4	Caractérisation des performances dans un milieu contrôlé . . . . .	42
2.5	Discussion à propos de la distribution spatiale des sources . . . . .	52
<b>3</b>	<b>Applications au SHM</b>	<b>63</b>
3.1	Estimation à partir de signaux quantifiés . . . . .	65
3.2	Applications de la modalité acoustique pour SHM . . . . .	66
3.3	Ouvertures vers un identificateur passif . . . . .	76
<b>4</b>	<b>Estimation bayésienne de la corrélation de Green</b>	<b>83</b>
4.1	Modèle pour la corrélation de Green . . . . .	85
4.2	Modélisation bayésienne et algorithme d'estimation . . . . .	86
4.3	Raffinements algorithmiques et ouvertures . . . . .	101
	<b>Conclusions et perspectives</b>	<b>111</b>



# Introduction à l'identification passive

---

Ce premier chapitre décrit le contexte et le paysage scientifique autour de la thématique de mon doctorat. Je propose également un exemple introductif, qui présentera les enjeux et les objectifs de mon travail. Je n'aurai de cesse d'y revenir par la suite.

Ce doctorat s'est déroulé entre les deux unités de recherche que sont le Gipsa-lab, département traitement du signal et le CEA-Leti, département des systèmes intégrés. Une des concrétisations de leur co-activité est la thématique de recherche de mon doctorat, qui peut être formulée ainsi :

*« Identification d'un milieu de propagation linéaire dans un contexte passif d'écoute. »*

Les résultats rapportés dans ce manuscrit sont de l'ordre de la modélisation, de la caractérisation de performances d'estimateurs, du développement d'algorithmiques et de protocoles expérimentaux.

Bonne lecture.

## Sommaire

<b>1.1</b>	<b>Exemple introductif : l'oscillateur amorti . . . . .</b>	<b>3</b>
1.1.1	Formule des interférences et identité de Ward . . . . .	4
1.1.2	Illustration, estimateurs et performances . . . . .	5
1.1.3	Analyses, enjeux et problématiques . . . . .	6
<b>1.2</b>	<b>Historique, état de l'art et positionnements . . . . .</b>	<b>8</b>
1.2.1	Prémices . . . . .	8
1.2.2	Aboutissement . . . . .	9
1.2.3	Positionnement de la modalité acoustique . . . . .	10
1.2.4	Positionnement par rapport au SHM passif . . . . .	11
<b>1.3</b>	<b>Travaux présentés . . . . .</b>	<b>19</b>

---



## 1.1 Exemple introductif : l'oscillateur amorti

Débutons en douceur par un en-cas introductif, afin de pénétrer la problématique de *l'identification passive*. Tout d'abord, derrière ce terme abscons se cachent en fait deux notions simples à définir. *Identifier* un système signifie estimer des paramètres qui le caractérisent. Le contexte *passif* dans lequel cette identification a lieu indique que l'opérateur ne maîtrise pas les sources qui viennent stimuler le système. Ces sources sont appelées sources d'opportunité et tout l'enjeu consiste à trouver comment les exploiter pour parvenir à identifier le système.

L'exemple illustre la majorité des enjeux liés à ce contexte d'estimation, et les travaux présentés dans la suite du manuscrit s'efforceront de les caractériser pour un milieu de propagation. Dans cet exemple, le système étudié n'est pas un milieu de propagation, mais un oscillateur amorti, ce compagnon de longue date pour le scientifique, qui a notamment permis à Planck de modéliser la matière, la reformulation des équations de Maxwell en faisant apparaître la notion de photon, ou encore à modéliser les vibrations élastiques de cristaux en prenant le nom de phonon.

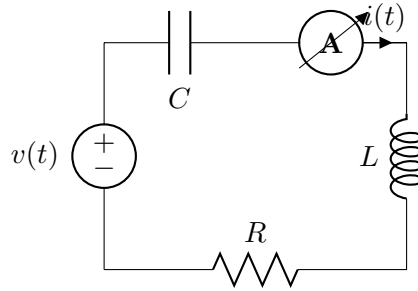


FIGURE 1.1 – Circuit RLC, oscillateur amorti

Représenté Fig. 1.1, l'oscillateur amorti voit sa charge de courant  $q(t)$  obéir à une équation différentielle du second degré en temps, qui s'obtient par exemple en considérant la loi des mailles sur l'unique boucle du circuit :

$$\ddot{q}(t) + 2\eta\dot{q}(t) + \omega_0^2 q(t) = \frac{1}{L}v(t) \quad (1.1)$$

où  $\omega_0 = \sqrt{1/LC}$  est la pulsation propre du circuit à laquelle la charge devrait osciller en l'absence d'amortissement, ce dernier étant modélisé par la constante  $\eta = R/2L$  (pertes par effet Joule). La réponse d'un système à une impulsion  $\delta(t)$  (exprimé en Volts) est sa réponse à la sollicitation la plus élémentaire qui soit et suffit à décrire exhaustivement le système. Pour l'oscillateur amorti, elle est notée  $h(t)$  et a pour forme :

$$h(t) = \frac{1}{L\omega} \exp(-\eta t) \sin(\omega t) u(t). \quad (1.2)$$

Dans l'Eq. (1.2),  $u(t)$  est l'échelon de Heaviside, tenant compte du fait que  $h(t)$  est une fonction causale. L'amortissement se traduit par le terme monotone décroissant  $\exp(-\eta t)$  et se retrouve également dans la partie oscillante de la réponse, en modifiant la pulsation d'oscillation  $\omega_0$  du système non amorti par une pulsation plus faible  $\omega = \sqrt{\omega_0^2 - \eta^2}$ .

L'objectif de cet exemple est le suivant : trouver un moyen d'estimer  $h(t)$  lorsque l'oscillateur est excité par une tension  $v(t)$  non contrôlée. Pour ce faire, l'approche est découpée en deux étapes et tous les calculs détaillés peuvent être trouvés dans l'Annexe A.

### 1.1.1 Formule des interférences et identité de Ward

Définissons la fonction  $C_f(t)$  qui à toute fonction réelle  $f$  associe une approximation de la fonction d'auto-corrélation de  $f$  si c'était une variable aléatoire stationnaire et ergodique :

$$C_f(t) \triangleq \int_{\mathbb{R}} f(t')f(t'+t)dt'. \quad (1.3)$$

Tous calculs faits, la corrélation de  $h(t)$  pour l'oscillateur amorti s'écrit :

$$C_h(t) = \frac{1}{4L^2\omega_0^2} \exp(-\eta|t|) \times \left[ \frac{\cos(\omega t)}{\eta} + \frac{\sin(\omega|t|)}{\omega} \right]. \quad (1.4)$$

Le terme d'amortissement ainsi que la pulsation d'oscillation se retrouvent dans la corrélation du filtre. La dérivée première de  $C_h(t)$  s'écrit :

$$\dot{C}_h(t) = -\frac{1}{4L^2\omega_0^2} \exp(-\eta|t|) \times \left[ \frac{\eta}{\omega} + \frac{\omega}{\eta} \right] \sin(\omega t), \quad (1.5)$$

et fait ainsi apparaître la relation cherchée entre  $h(t)$  et  $C_h(t)$ , appelée identité de Ward :

$$h(t) = -4\eta L \dot{C}_h(t) \times u(t). \quad (1.6)$$

Trouver l'équivalent de cette relation pour un milieu de propagation visco-acoustique sera fondamental pour formaliser le contexte d'estimation passive. En effet, nous montrerons que dans ce contexte il sera possible d'estimer l'équivalent de la fonction  $C_h(t)$  à partir des champs mesurés, et non pas la réponse impulsionnelle. L'application de l'identité de Ward visco-acoustique permettra alors d'estimer la réponse impulsionnelle dans un contexte passif d'estimation.

La seconde étape consiste à traduire la réponse du système à une excitation non contrôlée  $v(t)$ . Pour cela, notons que l'oscillateur agit comme un filtre sur l'excitation de telle sorte que sa réponse étant notée  $q(t)$  s'écrive :

$$q(t) \triangleq \int_{\mathbb{R}} h(t-t')v(t')dt' = (h * v)(t). \quad (1.7)$$

Supposons maintenant que la source  $v(t)$  soit un processus aléatoire stationnaire, de fonction d'auto-corrélation  $C_v(t)$ . La corrélation de l'évolution de la charge  $q(t)$  est obtenue en écrivant la formule des interférences dans le domaine temporel, soit :

$$C_q(t) = [C_v * C_h](t). \quad (1.8)$$

Ainsi, puisque les opérations d'intégration et de filtrage peuvent commuter, nous arrivons à étendre l'identité Eq. (1.6) de la manière suivante :

$$(C_v * h)(t) \times u(t) = -4\eta L \dot{C}_q(t) \times u(t). \quad (1.9)$$

L'accès par la mesure à  $C_h(t)$  est tributaire des propriétés spectrales de la source. Si la source  $v(t)$  est stationnaire blanche,  $C_v(t)$  vaut  $P_0\delta_t$ , élément neutre du produit de convolution. Alors,  $C_q(t)$  est proportionnel à  $C_h(t)$ , auquel cas l'identité Eq. (1.6) n'est plus uniquement vérifiée pour une source impulsionnelle. Nous avons donc obtenu une relation permettant de retrouver  $h(t)$  dans un cas particulier où la source excitatrice d'opportunité du système est un bruit stationnaire. La reconstruction peut être optimale si la corrélation de la source est connue, ce qui est improbable dans le contexte passif, ou bien si la source est blanche.

### 1.1.2 Illustration, estimateurs et performances

La Fig. 1.2 illustre l'exploitation pratique de l'identité obtenue Eq. (1.6). La réponse impulsionnelle de l'oscillateur amorti est  $h(t)$  et ce dernier est excité par la source  $v(t)$ , variable aléatoire Gaussienne centrée et blanche. La réponse de l'oscillateur est  $q(t)$ . Dans un cadre d'acquisition réaliste, la mesure  $m[n]$  est l'échantillonnage-quantification de la charge vraie, bruitée par l'instrumentation, modélisé par une séquence i.i.d. Gaussienne centrée, et telle que le RSB vaut 2 dB.

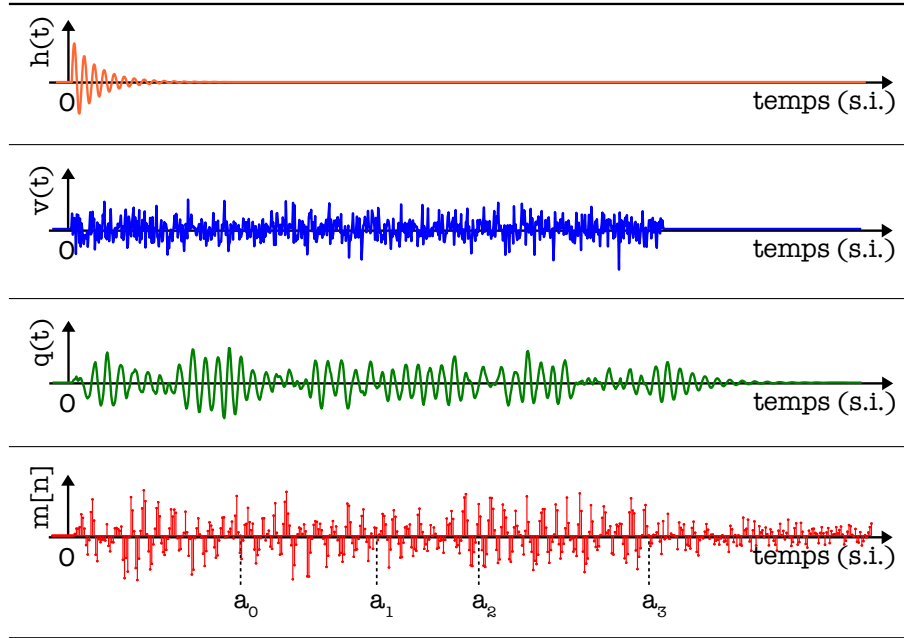


FIGURE 1.2 – Excitation d'un oscillateur amorti dans un contexte passif.

Pour retrouver  $h(t)$ , il faut lui construire un estimateur, c'est-à-dire transformer judicieusement la mesure  $m[n]$  pour reconstruire la fonction  $h(t)$ . Ici, pour exploiter l'identité (1.6), il s'agira d'abord d'estimer la corrélation de la charge mesurée. La source étant stationnaire



et le système linéaire, la réponse  $m[n]$  est également une séquence stationnaire (en dehors des effets de bord du filtrage) et ergodique. Ainsi, de manière classique, un estimateur pour  $C_h[n]$  peut être construit en calculant l'espérance temporelle et spatiale des champs numérisés à l'aide d'une fenêtre glissante d'analyse. La corrélation estimée est la moyenne d'ensemble des corrélations estimées sur chacune des  $K$  fenêtres d'analyse :

$$\hat{C}_h[n] = \frac{1}{K} \sum_{k=1}^K \frac{1}{(b_k - a_k)} \sum_{l=a_k}^{b_k-n} m[l]m[l-n], \quad (1.10)$$

où les  $a_k$  et  $b_k$  définissent les extrémités de la fenêtre.

Cet estimateur est caractérisé par un ensemble de performances étudiées dans la littérature. Il est biaisé mais à une variance faible, voir [1]. Le biais, convolutif (fenêtre de Barlett), est compensable par ailleurs par l'utilisation d'une fenêtre d'apodisation et son choix dépend de l'importance apportée à la déformation induite par les lobes secondaires. La connaissance et la maîtrise des performances des estimateurs est cruciale, puisqu'elles conditionnent la bonne exploitation des résultats. En particulier ici,  $\hat{C}_h[n]$  n'est pas la finalité, puisque c'est un intermédiaire pour accéder à  $\hat{h}[n]$ . Des exemples des résultats d'estimation sont illustrés dans la Fig. 1.3, où 1, 10, 100 puis 1000 segments extraits de la réponse  $m[n]$  sont utilisés pour calculer  $\hat{C}_h[n]$  ainsi que  $\hat{h}[n]$  par application de l'identité (1.6).

Une observation strictement qualitative permet de rendre compte d'ores et déjà que les performances d'estimation de chacun des paramètres n'est pas identique sur les deux fonctionnelles. En particulier, la pulsation reconstruite semble davantage fidèle à la pulsation vraie que l'amortissement. Son estimation semble d'ailleurs meilleure sur  $\hat{C}_h[n]$  que sur  $\hat{h}[n]$ . Cette remarque tient du fait que bien que l'identité (1.6) conduise à l'estimateur de  $\hat{h}$  qui contient l'information sur l'amortissement, son accès est tributaire de la propagation des incertitudes induite par la cascade d'estimations (accès à  $\hat{C}_h[n]$  puis application de l'identité (1.6)).

### 1.1.3 Analyses, enjeux et problématiques

De l'exemple mis en oeuvre, tirons des conclusions.

Premièrement, récapitulons que l'Eq. (1.6) offre un moyen de reconstruire la réponse impulsionnelle d'un système, et donc par extension, de paramètres qui en découlent, à partir de la corrélation de la réponse du système à une excitation aléatoire stationnaire. Cette relation permet de reconstruire précisément  $h(t)$ , si la corrélation de la source est connue, en appliquant un algorithme de déconvolution. Cette configuration est cependant improbable dans le contexte passif, où la source est peu maîtrisée et connue. A défaut, la reconstruction peut encore être optimale si la source est blanche.

En situation réelle, la réponse  $m(t)$  est mesurée, impliquant une étape d'échantillonnage et de quantification. La conséquence de cette numérisation, couplée à l'effet de la chaîne d'acquisition, est que la mesure est une version bruitée et parfois distordue de la grandeur physique mesurée. A cela s'ajoute le fait que les quantités d'intérêt sont reconstruites par l'intermédiaire

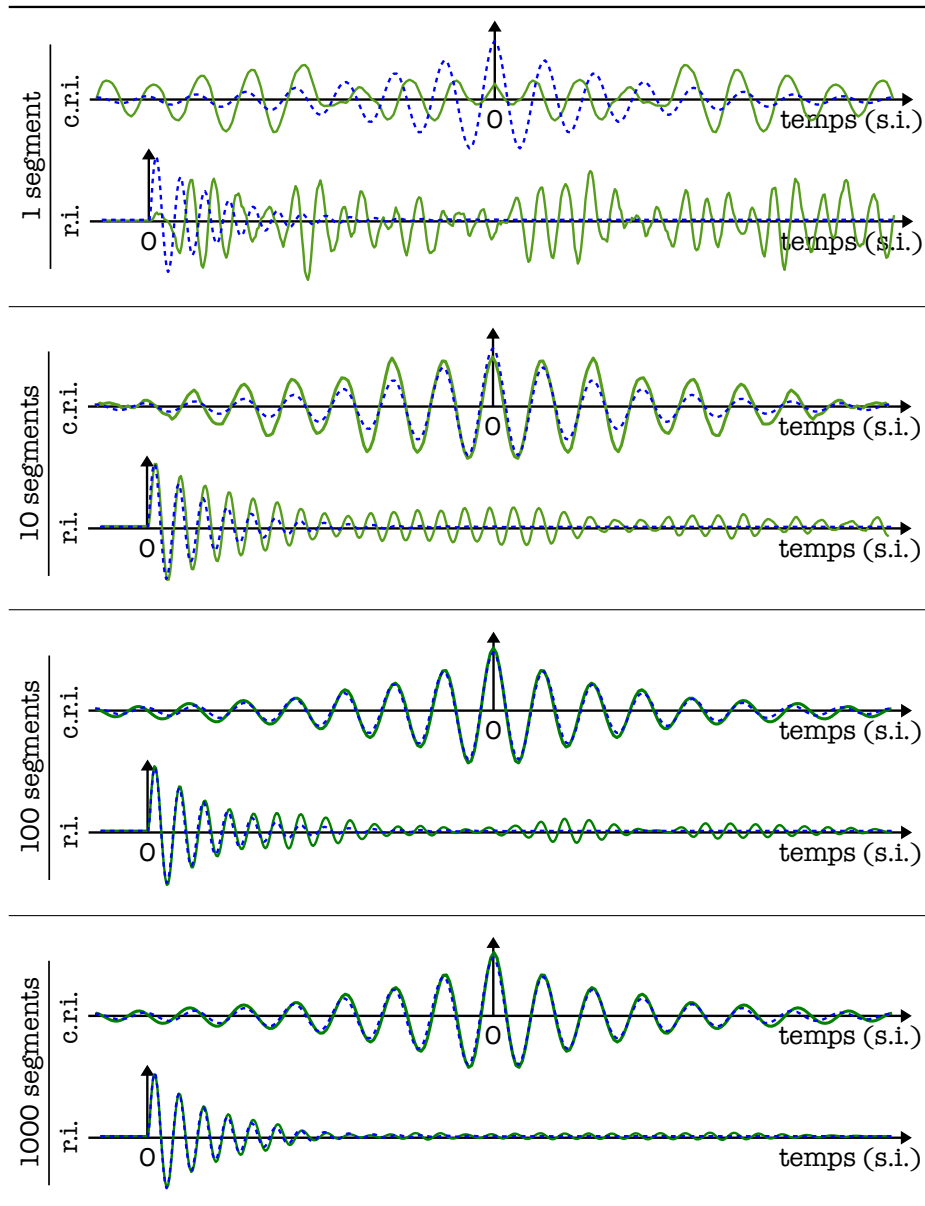


FIGURE 1.3 – Estimation de la corrélation de la réponse impulsionnelle (c.r.i.) et de la réponse impulsionnelle (r.i.) avec  $\{1, 10, 100, 1000\}$  segments extraits de la réponse  $m[n]$ . (Bleu)  $C_q(t)$  théorique. (Vert) Fonctions estimées.

d'estimateurs. Ces derniers possèdent des performances statistiques et de convergence, ainsi qu'une certaine robustesse à l'écart aux hypothèses faites dans le modèle d'observation. Dès lors, les estimations successives comme avec l'application de l'identité Eq. (1.6) propagent les incertitudes déjà présentes à la mesure. Caractériser cette propagation d'erreur est un pas important vers la maîtrise du protocole d'estimation en contexte passif.

Enfin, l'identité (1.6) est construite pour un système régi par une équation différentielle en temps. Son extension pour un milieu régi par une équation de propagation n'est pas complète-

ment transparent. La dissipation en particulier apparaît comme une composante primordiale pour faire émerger l'Eq. (1.6). Obtenue ici pour la modélisation de frottements fluides, que se passerait-il si le système n'était pas amorti ? Que se passerait-il pour d'autres modèles de dissipation, plus à même de décrire la propagation dans certains milieux ? Notons finalement par l'extension à un système évoluant dans le temps et dans l'espace, l'intérêt que représente l'étude du rôle des propriétés spectrales en temps et en espace du champ mesuré sur les estimations en contexte passif.

## 1.2 Historique, état de l'art et positionnements

En théorie de l'identification des systèmes linéaires, les moments statistiques sont des outils appréciés et qui ont été très largement exploitées depuis plusieurs décennies. Le théorème des interférences offre une première relation entre les corrélations de la réponse impulsionnelle du système, du champ mesuré et de la source. Le filtrage adapté est une autre technique qui exploite également la corrélation du signal source et qui trouve historiquement des applications avec le radar où la modalité consiste à estimer par exemple une distance de propagation. Certaines approches ont émergé pour traiter des classes de signaux non stationnaires, telle que la cyclostationnarité (statistiques cycliques), dans le domaine de *l'identification aveugle de canal* en télécommunication, ou encore pour le contrôle de l'état de santé de machines tournantes à l'aide de signaux vibratoires. Avec l'aperçu apporté par l'oscillateur amorti, nous avons compris que pour le contexte passif d'estimation, c'est la physique qui va venir offrir la relation qui permettra d'identifier le système à partir de la corrélation des champs mesurés. Dans ce qui suit, un rapide état de l'art de cette thématique est exposé.

### 1.2.1 Prémisses

Dans ses travaux pionniers sur l'agitation thermique dans une résistance [2], Nyquist formalise un phénomène physique observé expérimentalement par J. B. Johnson. Il répond à ce dernier dans une lettre de 2 pages, d'une concision exemplaire, en énonçant une relation qui sera reprise quelques décennies plus tard pour construire un théorème, appelé « théorème fluctuation-dissipation ». Pour consulter un ouvrage étendu sur le sujet, se référer par exemple au travail de Kubo [3]. Dans ce théorème, établi par H. B. Callen et T. A. Welton, deux quantités issues d'un système dissipatif linéaire sont mises en regard : les fluctuations de la réponse du système, représentées par la corrélation du champ mesuré et la dissipation, partie imaginaire de la réponse du système dans le domaine des fréquences. A l'époque, ce théorème n'était formalisé que pour traiter le cas des fluctuations thermiques et quantiques de systèmes linéaires. Aperçu sous la forme de l'Eq. (1.6) pour l'oscillateur amorti, il a depuis largement étendu sa portée, grâce à divers rapprochements qui ont été faits avec la propagation des ondes dans les milieux linéaires.

Cette extension débute en 1968, alors que Claerbout ?? travaille dans le domaine de l'exploration géophysique où l'objectif est de caractériser les propriétés physiques de l'intérieur de

la Terre. En considérant que cet intérieur est construit selon un modèle en couches horizontales, une source d'onde plane à la surface libre verrait le front d'onde traverser les couches de plus en plus profondes, se séparant en ondes réfléchies et transmises à chaque interface. L'outil au coeur des travaux de ce manuscrit, la corrélation, est utilisé à la manière du radar, pour mesurer les instants de réflexion et donc par extension, la profondeur des interfaces. Cette approche a permis de mettre en évidence le fait que la corrélation du champ mesuré correspondait au filtrage de la fonction de corrélation de la source par un milieu virtuel, dérivé du milieu original. De son expérience, Claerbout nourrit une intuition, formulée clairement dans ses travaux 30 ans plus tard. Il travaille alors dans le domaine de l'héliosismologie, qui consiste à étudier l'activité sismique du soleil par son observation depuis la Terre. Il présume ainsi :

*« By cross-correlating noise traces recorded at two locations on the surface, we can construct the wavefield that would be recorded at one of the locations if there was a source at the other. »*

Cette conjecture, reprenant mot pour mot le concept derrière le théorème fluctuation-dissipation, tarde à être vérifiée expérimentalement, faute de données exploitables. C'est dans sa communication avec Rickett, voir [4], qu'il diffuse enfin la confirmation expérimentale de son intuition. En parlant de *noise traces*, le lecteur fera le lien avec la source choisie stationnaire et blanche qui est venue exciter l'oscillateur amorti de l'exemple 1.1. Le pendant de ce type de sources pour des ondes de propagation sismique est étudié par Aki [5] et date des années 70. Il s'intéresse plus particulièrement aux champs réverbérés créés par des événements sismiques, qu'il appelle *coda waves*, en référence au symbole musical qui indique la conclusion du morceau et qui invite à se concentrer sur la mort de la ligne harmonique finale. En caractérisant ses propriétés d'atténuation et son contenu spectral, il donne naissance à une discipline aujourd'hui connue sous le nom de l'interférométrie sismique.

### 1.2.2 Aboutissement

Au début des années 2000, Campillo et Paul [6] suggèrent, fort d'un recul sur les résultats précédents et d'une intuition remarquable, une méthode afin de combiner les travaux de Aki et ceux de Claerbout. Ils exposent ainsi à la communauté scientifique un protocole pour pratiquer l'interférométrie à partir de codas sismiques. Validée sur des données expérimentales du Mexique, l'approche expérimentale est décrite avec suffisamment de précision pour être reproduite et étudiée.

En salle d'expérimentation, ce sont Weaver et Lobkis qui formalisent les premiers cette avancée, en s'appuyant sur une expérience répétée et fiable de propagation ultrasonique dans un bloc d'aluminium, voir [7, 8]. Leur approche, qui consiste à modéliser un champ diffus dans un corps fini par une représentation modale adaptée, mène sur un calcul de la corrélation du champ diffus. Moyennant des hypothèses supplémentaires de blancheur et d'équipartition de l'énergie dans les modes, ils parviennent à formaliser une relation fonctionnelle entre la corrélation du bruit et la réponse impulsionnelle du milieu. Cette relation, discutée, reprise et corrigée plus tard par la communauté a véritablement permis de poser une base solide de

réflexion autour des choix algorithmes et des justifications physiques permettant d'expliquer un résultat d'expérience robuste et répétable et jusqu'alors fort intrigant.

Durant la décennie suivante, la formalisation se poursuit en regard de la multiplication des expériences réalisées. En sismologie, des ouvertures à différents types d'ondes, présentes à différentes échelles sont proposées par des chercheurs comme Roux [9], Prieto [10], Shapiro [11], Stehly [12] et d'autres [13, 14]. De la ville au continent, et plus récemment à l'échelle de la planète entière, l'interférométrie sismique passive est la discipline qui témoigne le plus de la maîtrise grandissante de la méthode et du savoir-faire à exploiter le bruit ambiant de la Terre. Idéale à mettre en oeuvre à l'échelle du laboratoire, l'extension de l'approche aux milieux hétérogènes est étudiée notamment par Larose [15, 16] et Levine [17]. Le succès de la robustesse de l'approche sur données réelles incite à explorer de nouvelles modalités et très vite, la communauté s'attaque à l'acoustique aérienne et sous-marine, [18, 19], aux ondes électro-magnétique [20], à la tomographie de la Lune [21] et d'autres applications tout aussi originales, voir [22, 23, 24, 25]. En parallèle, alors que la méthodologie est continuellement validée par l'expérience, se montrant robuste et répétable dans divers contextes très variés, des travaux d'unification de la théorie derrière ce phénomène sont menés. Nous noterons un tutoriel en deux parties, autant théorique que pratique de la part de Wapenaar [26, 27]. De son côté, Colin de Verdière apporte une contribution significative en approchant le problème avec le formalisme des opérateurs [28, 29, 30]. Des travaux multidisciplinaires regroupent et analysent des débats bien avancés, principalement autour de justifications sur la modélisation de la physique, voir les travaux de Gouédard [31], de Larose [32], de Snieder [33, 34], de Roux [35] et d'autres [36, 37, 38]. Un article de synthèse est proposé par les acteurs majeurs de l'estimation passive, Weaver, Campillo et Larose [39] et un livre regroupant les interventions majeures sur le sujet et ses perspectives est édité par Wright et Weaver [40].

A la fin des années 2000, les travaux d'unification s'arrêtent et laissent en suspens certains débats. Malgré tout, la communauté scientifique dispose de réponses suffisamment éclairées sur la méthodologie de l'identification à partir de sources d'opportunité. Le temps de l'application est venu, et c'est maintenant le raffinement du protocole qui est convoité, et les performances d'estimation sont à présent scrupuleusement recherchées. Parmi les travaux d'ouverture, la thématique d'auto-localisation d'un réseau de capteurs à partir de bruit ambiant est évoquée par Sabra [41], alors que de manière connexe la stabilité et la détection de la dérive de l'instrumentation est diagnostiquée dans ce même contexte par Stehly [42]. La course à la réduction des données à traiter est abordée dans [43, 44]. Afin d'améliorer la reconstruction des fonctions de Green, des algorithmes itératifs sont également proposés, voir [45, 46]. Enfin, notons l'engouement pour la fusion de modalités, destinée principalement à améliorer la résolution des grandeurs reconstruites, voir par exemple [47, 48].

### 1.2.3 Positionnement de la modalité acoustique

Au CEA-Leti DSIS, les technologies en rupture ont le vent en poupe et un des moteurs principaux de l'activité scientifique du département est la création de systèmes innovants. Son savoir-faire trouve une véritable force dans sa capacité à manipuler une chaîne de mesure, de

sa conception à son exploitation dans des environnements industriels contraints. La thèse de Mikael Carmona [19], soutenue en 2011, est en accord avec cette expertise et s'inscrit dans le cadre de l'estimation à partir de sources d'opportunité. Dans ses travaux, il est amené à s'intéresser à la modalité acoustique aérienne, voir par exemple l'article [18], pour sa facilité de mise en oeuvre. Perçu d'abord comme un cas d'école, le champ acoustique répond à l'équation scalaire de Helmholtz, bien plus facile à manipuler que la propagation d'ondes sismiques. C'est en quelque sorte le pendant le plus simple qui soit de l'oscillateur amorti pour un milieu de propagation. La simulation par propagation de rayons ou par représentation modale sont des outils faciles à mettre en oeuvre, rendant son étude théorique aisée et les longueurs d'ondes manipulées permettent par ailleurs l'expérimentation dans des salles de laboratoires classiques, avec du matériel de mesure acoustique standard. Notons par ailleurs que les champs acoustiques réverbérés sont intensivement étudiés par les acousticiens, même si la finalité est bien différente de la nôtre.

Ce manuscrit se concentre en réalité presque exclusivement sur cette modalité acoustique qui, loin d'être un cas d'école, trouvera des applications pertinentes. Etant donnée la dynamique récemment empruntée par la communauté, il apparaît opportun de faire entrer en jeu les outils du traitement du signal, plus qu'indiqués pour modéliser et caractériser des performances d'estimateurs. Ainsi, l'objectif de ce manuscrit est clair, il s'agit de développer une théorie de l'estimation et détection dans le contexte passif, appliquée à l'acoustique aérienne. Le point de départ consistait à capitaliser sur les travaux de Mikael Carmona, en particulier en reprenant l'article [18] dans lequel il formalise l'identité Eq. (1.6) tant convoitée pour les milieux visco-acoustiques, et dans lequel il expose un ensemble de résultats et d'analyses qui ne nous laissent pas vierge sur le sujet. Du point de vue expérimental, la chaîne d'acquisition était à portée de la main, autorisant très rapidement à éprouver l'algorithme d'identification passive et à le travailler.

#### 1.2.4 Positionnement par rapport au SHM passif

SHM signifie en anglais *Structural Health Monitoring*. L'état de santé d'une structure est une mesure, par l'intermédiaire d'indicateurs, de sa capacité à effectuer sa fonction. Cette capacité est dégradée par le vieillissement des matériaux, soit naturellement, soit à la suite d'accidents. Chaque sollicitation mécanique accélère ce vieillissement et un incident ou des conditions climatiques extrêmes peuvent aggraver brutalement l'état de santé de la structure, voire même la mettre hors d'usage [49]. Les conséquences, principalement financières (et pas des moindres), sont également humaines. Des approches sont étudiées depuis les années 60 afin de mesurer l'évolution des caractéristiques physiques et géométriques des structures, voir par exemple [50, 51] et une thèse complète sur le sujet [52]. Les méthodes de SHM peuvent être actives, par exemple par la mesure d'une réponse impulsionnelle, lorsque celle-ci est générée par l'explosion d'une charge, mais les approches passives sont préférées, puisque leur impact sur la structure est moindre, voir par exemple [53, 54]. Tout le challenge consiste alors à obtenir des résultats exploitables de manière équivalente. D'une certaine manière, l'état de l'art en matière d'interférométrie sismique est également celui du SHM passif et certains travaux, comme ceux de Cole [55, 56], cherchaient déjà à exploiter les vibrations inhérentes aux structures en cours

d'utilisation pour les identifier. En l'occurrence, Cole cherchait à caractériser les résonances de pièces de moteurs de fusées pour la NASA, sans employer de transformées rapides (FFT), qui n'étaient pas encore popularisées.

Le pont, la route, l'immeuble et le rail sont les quatre types de structures qui décrivent la majorité des milieux sujets au SHM en génie civil, tandis que dans l'aéronautique, ce sont principalement les ailes d'avion qui sont écoutées. Le SHM est une discipline à parfait mi-chemin entre la recherche et l'industrie. Les acteurs académiques sont majoritairement italiens, asiatiques et américains tandis que le monde industriel est davantage représenté par la France et les Etats-Unis. Les grands comptes ont des besoins et des start-up apportent des réponses à ces besoins. Au centre, le monde académique fait l'interface, créant la rupture technologique et méthodologique. Les technologies employées sont très variées, allant de la fibre optique aux piezo pour la mesure de contraintes, et en passant par la recherche de matériaux composites multi-modaux en aéronautique. L'utilisation intensive d'accéléromètres et de lasers pour la mesure d'accélération semble être quant à elle la modalité préférée entre toutes.

Une technique classique exploitant cette mesure d'accélération est l'analyse modale des structures. Ces dernières, excitées par diverses sources de natures variées ont une réponse à ces sollicitations qui contiennent de l'information sur la structure elle-même. En particulier, la densité spectrale de puissance est une image (surjective) de la géométrie de la structure. En guise d'illustration, observons le spectre basses fréquences ( $< 10$  Hz) d'une structure emblématique au CEA-Leti : la liaison qui relie deux salles blanches (un pont sur lequel circule une nacelle électrique). Pour ce faire, nous avons accroché un accéléromètre à un pilier de la structure et enregistré la réponse de la structure aux sollicitations extérieures pendant 20 min. Un algorithme d'estimation de spectre de type Welch est alors employé sur ce « bruit ambiant », voir Fig. 1.5. Des modes apparaissent, liés à la géométrie de la structure. Si cette dernière venait à se déformer, le spectre estimé changerait également et une stratégie de suivi de l'état de santé du bâtiment construite sur ces changements permettrait d'en évaluer la gravité.



FIGURE 1.4 – Liaison Blanc-Blanc du Leti.

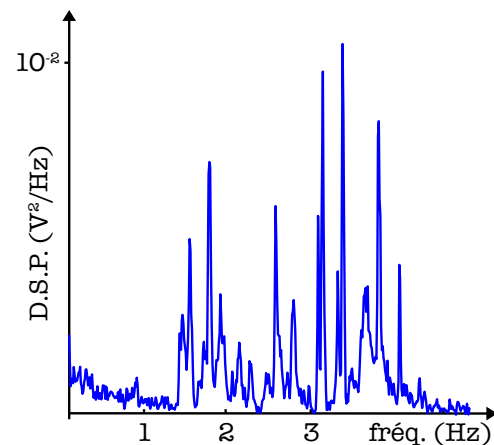


FIGURE 1.5 – Spectre de Welch du bruit.

Ce que je retiens de mon aperçu de la communauté SHM au cours des conférences auxquelles j'ai assistées, c'est que les spécificités des structures freinent considérablement la nor-



malisation des approches et l'établissement de modèles génériques. Pour illustrer ce propos, imaginez que la tour la plus haute du monde, à Dubaï, n'est instrumentée que par trois accéléromètres, tandis que le pont de Sydney est recouvert de plus de 1000 capteurs de toutes sortes.

Dans le travail exposé dans le manuscrit, la modalité acoustique sera confrontée au monde du SHM, extension assez naturelle de l'état de l'art exposé jusqu'à présent. Il se trouve que ce choix n'est pas sans conséquences, puisque cette modalité est complémentaire des techniques existantes en SHM sur plusieurs points. Contrairement aux ondes sismiques, les longueurs d'ondes en jeu avec l'acoustique ouvrent la porte à l'estimation de paramètres géométriques avec une résolution excellente, autorisant par exemple les techniques de localisation de réseaux de capteurs et de sources. De plus, la faible dissipation à ces longueurs d'ondes étend la portée de l'approche à plusieurs dizaines de mètres, contrairement aux techniques de détection d'hétérogénéité. L'équivalent de la détection d'hétérogénéité par des ultrasons, en acoustique aérienne, correspond à la détection d'objets de quelques centimètres d'envergure. Enfin, l'estimation de paramètres modaux n'est pas uniquement accessible à la modalité sismique ; c'est tout un pan du traitement acoustique des salles qui utilise cet outil, à des fins de confort d'écoute musicale par exemple (jusqu'à 500 Hz environ).

L'angle d'approche choisi pour apporter la modalité acoustique au SHM commence par la caractérisation des performances d'estimation des paramètres. C'est un point d'entrée qui permettra d'améliorer le protocole expérimental et d'obtenir un retour quantifié impartial sur les conditions expérimentales rencontrées. Le développement du SHM passif acoustique sera alors un support d'illustration et une extension intéressée de la théorie de l'estimation passive développée dans le manuscrit.





# Références

- [1] H. L. VAN TREES, K. L. BELL et Z. TIAN, *Detection, estimation and filtering theory*, WILEY, éd. 1968 (cf. p. 6).
- [2] H. NYQUIST, “Thermal agitation of electric charge in conductors”, *Physical Review E*, t. 32, n° 1918, p. 110–113, 1927 (cf. p. 8).
- [3] R. KUBO, “The fluctuation-dissipation theorem”, rap. tech. (cf. p. 8).
- [4] J. RICKETT et J. CLAERBOUT, “Acoustic daylight imaging via spectral factorization : Helioseismology and reservoir monitoring”, rap. tech., 1999, p. 171–181 (cf. p. 9).
- [5] K. AKI et B. CHOUET, “Origin of Coda Waves : source, attenuation, and scattering effects”, *Journal of Geophysical Research*, t. 80, n° 23, p. 3322–3342, 1975 (cf. p. 9).
- [6] M. CAMPILLO et A. PAUL, “Long-Range Correlations in the Diffuse Seismic Coda”, *Nature Reports*, t. 299, n° January, p. 547–549, 2003 (cf. p. 9).
- [7] O. I. LOBKIS et R. L. WEAVER, “On the emergence of the Green’s function in the correlations of a diffuse field”, *Acoustics* 40, t. 110, n° September, p. 3011–3017, 2001 (cf. p. 9).
- [8] R. WEAVER et O. LOBKIS, “On the emergence of the Green’s function in the correlations of a diffuse field : pulse-echo using thermal phonons”, *Ultrasonics*, t. 40, n° 1-8, p. 435–439, mai 2002 (cf. p. 9).
- [9] P. ROUX, “P-waves from cross-correlation of seismic noise”, *Geophysical Research Letters*, t. 32, n° 19, p. L19303, 2005 (cf. p. 10).
- [10] G. a. PRIETO, J. F. LAWRENCE et G. C. BEROZA, “Anelastic Earth structure from the coherency of the ambient seismic field”, *Journal of Geophysical Research*, t. 114, n° B7, B07303, juil. 2009 (cf. p. 10).
- [11] N. M. SHAPIRO et M. CAMPILLO, “Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise”, *Geophysical Research Letters*, t. 31, n° April, p. 8–11, 2004 (cf. p. 10).
- [12] L. STEHLY, B. FRY, M. CAMPILLO, N. M. SHAPIRO, J. GUILBERT, L. BOSCHI et D. GIARDINI, “Tomography of the Alpine region from observations of seismic ambient noise”, *Geophysical Journal International*, t. 178, p. 338–350, 2009 (cf. p. 10).
- [13] C. NUNZIATA, G. D. NISCO et G. F. PANZA, “S-waves profiles from noise cross correlation at small scale”, *Engineering Geology*, t. 105, n° 3-4, p. 161–170, 2009 (cf. p. 10).
- [14] P. POLI, H. A. PEDERSEN et M. CAMPILLO, “Emergence of body waves from cross-correlation of short period seismic noise”, *Geophysical Journal International*, t. 188, p. 549–558, 2012 (cf. p. 10).
- [15] E. LAROSE, “Diffusion multiple des ondes sismiques et expériences analogiques en ultrasons”, thèse de doct., Université Joseph Fourier, 2005 (cf. p. 10).
- [16] E. LAROSE, T. PLANES, V. ROSSETTO et L. MARGERIN, “Locating a small change in a multiple scattering environment”, *Applied Physics Letters*, t. 20, p. 1–9, 2010 (cf. p. 10).

- [17] A. J. LEVINE et T. C. LUBENSKY, “Two-point microrheology and the electrostatic analogy”, *Physical Review E*, t. 65, n° May, p. 1–13, 2001 (cf. p. 10).
- [18] M. CARMONA, O. J. J. MICHEL, N. SPRYNSKI, J.-L. LACOUME et B. NICOLAS, “Ward identities for visco-acoustic and visco-elastic propagation media”, *Wave Motion*, t. 49, n° 4, p. 484–489, 2012 (cf. p. 10, 11).
- [19] M. CARMONA, “Identification passive des milieux. Application à la reconstruction de géométries dans les réseaux de capteurs et au diagnostic des structures”, thèse de doct., Université de Grenoble, 2008 (cf. p. 10, 11).
- [20] E. SLOB et K. WAPENAAR, “Electromagnetic Green’s functions retrieval by cross-correlation and cross-convolution in media with losses”, *Geophysical Research Letters*, t. 34, n° 5, p. L05307, mar. 2007 (cf. p. 10).
- [21] C. SENS-SCHÖNFELDER et E. LAROSE, “Lunar noise correlation, imaging and monitoring”, eng, *Earthquake science*, t. 23, n° 5, p. 519–530, (cf. p. 10).
- [22] K. WAPENAAR, “Retrieving the Elastodynamic Green’s Function of an Arbitrary Inhomogeneous Medium by Cross Correlation”, *Physical Review Letters*, t. 93, n° 25, p. 254301, déc. 2004 (cf. p. 10).
- [23] G. MAINSANT, E. LAROSE, C. BRÖNNIMANN, D. JONGMANS, C. MICHOUUD et M. JABOYEDOFF, “Ambient seismic noise monitoring of a clay landslide : Toward failure prediction”, *Journal of Geophysical Research*, t. 117, p. 1–12, 2012 (cf. p. 10).
- [24] D. MARSAN, J. WEISS, E. LAROSE et J.-P. MÉTAXIAN, “Sea-ice thickness measurement based on the dispersion of ice swell”, *J. Acoust. Soc. Am.*, t. 131, n° 1, 2012 (cf. p. 10).
- [25] N. NAKATA, R. SNIEDER, T. TSUJI, K. LARNER et T. MATSUOKA, “Shear wave imaging from traffic noise using seismic interferometry by cross-coherence”, *Geophysics*, t. 76, n° 6, 2012 (cf. p. 10).
- [26] K. WAPENAAR, D. DRAGANOV, R. SNIEDER, X. CAMPMAN et A. VERDEL, “Tutorial on seismic interferometry : Part 1 — Basic principles and applications”, *Geophysics*, t. 75, n° 5, 75A195–75A209, sept. 2010 (cf. p. 10).
- [27] K. WAPENAAR, E. SLOB, R. SNIEDER et A. CURTIS, “Tutorial on seismic interferometry : Part 2 — Underlying theory and new advances”, *Geophysics*, t. 75, n° 5, 75A211–75A227, sept. 2010 (cf. p. 10).
- [28] Y. COLIN DE VERDIÈRE, “Semiclassical analysis and passive imaging”, *Nonlinearity*, t. 22, n° 6, R45–R75, juin 2009 (cf. p. 10).
- [29] Y. C. D. VERDIÈRES, “Mathematical models for passive imaging I : general background.”, p. 1–25, 2006 (cf. p. 10).
- [30] —, “Mathematical models for passive imaging II : Effective Hamiltonians associated to surface waves”, rap. tech., 2006, p. 1–9 (cf. p. 10).
- [31] P. GOUÉDARD, L. STEHLY, F. BRENGUIER, M. CAMPILLO, Y. C. D. VERDI, E. LAROSE, L. MARGERIN, P. ROUX et R. L. WEAVER, “Cross-correlation of random fields : mathematical approach and applications”, *Geophysical Prospecting*, n° 56, p. 375–393, 2008 (cf. p. 10).
- [32] E. LAROSE, L. MARGERIN, A. DERODE, B. van TIGGELEN, M. CAMPILLO, N. SHAPIRO, A. PAUL, L. STEHLY et M. TANTER, “Correlation of random wavefields : An interdisciplinary review”, *Geophysics*, t. 71, n° 4, SI11–SI21, juil. 2006 (cf. p. 10).

- [33] R. SNIEDER, Y. FAN, E. SLOB et K. WAPENAAR, “Equipartitioning is not sufficient for Green’s function extraction”, *CWP-600*, (cf. p. 10).
- [34] R. SNIEDER, “Retrieving the Green’s function of the diffusion equation from the response to a random forcing”, *Physical Review E*, t. 74, n° 4, p. 046 620, oct. 2006 (cf. p. 10).
- [35] P. ROUX, K. G. SABRA, W. a. KUPERMAN et A. ROUX, “Ambient noise cross correlation in free space : Theoretical approach”, *The Journal of the Acoustical Society of America*, t. 117, n° 1, p. 79, 2005 (cf. p. 10).
- [36] K. WAPENAAR, E. SLOB et R. SNIEDER, “Unified Green’s function retrieval by cross-correlation”, *CWP*, t. 571, 2004 (cf. p. 10).
- [37] R. L. WEAVER, “Ward identities and the retrieval of Green’s functions in the correlations of a diffuse field”, *Wave Motion*, t. 45, p. 596–604, 2007 (cf. p. 10).
- [38] R. SNIEDER, K. WAPENAAR et U. WEGLER, “Unified Green’s function retrieval by cross-correlation ; connection with energy principles”, *Physical Review E*, t. 75, n° 3, p. 036 103, mar. 2007 (cf. p. 10).
- [39] R. L. WEAVER, E. LAROSE et M. CAMPILLO, “On the precision of noise correlation interferometry”, *Geophysical Journal International*, t. 185, p. 1384–1392, 2011 (cf. p. 10).
- [40] M. WRIGHT et R. WEAVER, *New directions in linear acoustics and vibration*. New York : Cambridge University Press, 2010 (cf. p. 10).
- [41] K. SABRA, P. ROUX, a.M. THODE, G. D’SPAIN, W. HODGKISS et W. KUPERMAN, “Using Ocean Ambient Noise for Array Self-Localization and Self-Synchronization”, *IEEE Journal of Oceanic Engineering*, t. 30, n° 2, p. 338–347, avr. 2005 (cf. p. 10).
- [42] L. STEHLY, M. CAMPILLO et N. M. SHAPIRO, “Traveltime measurements from noise correlation : stability and detection of instrumental time-shifts”, *Geophysical Prospecting*, n° 171, p. 223–230, 2007 (cf. p. 10).
- [43] P. CUPILLARD, L. STEHLY et B. ROMANOWICZ, “The one-bit noise correlation : a theory based on the concepts of coherent and incoherent noise”, *Geophysical Journal International*, t. 184, p. 1397–1414, 2011 (cf. p. 10).
- [44] S. HIRATA, M. K. KUROSAWA et T. KATAGIRI, “Cross-Correlation by Single-bit Signal Processing for Ultrasonic”, *IEICE Trans. Fundamentals*, t. E91, n° 4, p. 1031–1037, 2008 (cf. p. 10).
- [45] B. FROMENT, M. CAMPILLO et P. ROUX, “Reconstructing the Green’s function through iteration of correlations”, *C. R. Geoscience*, t. 343, n° 8-9, p. 623–632, 2011 (cf. p. 10).
- [46] L. STEHLY, M. CAMPILLO, B. FROMENT et R. L. WEAVER, “Reconstructing Green’s function by correlation of the coda of the correlation (C3) of ambient seismic noise”, *Journal of Geophysical Research*, t. 113, p. 1–10, 2008 (cf. p. 10).
- [47] C. HADZIOANNOU, E. LAROSE, A. BAIG, P. ROUX et M. CAMPILLO, “Improving temporal resolution in ambient noise monitoring of seismic wave speed”, *Journal of Geophysical Research*, t. 116, p. 1–10, 2011 (cf. p. 10).
- [48] L. STEHLY, P. CUPILLARD et B. ROMANOWICZ, “Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise”, *International Geophysics*, t. 343, p. 591–599, 2011 (cf. p. 10).

- [49] N. NAKATA et R. SNIEDER, “Near surface weakening in Japan after the 2011 Tohoku Oki earthquake”, *Geophysical Research Letters*, t. 38, n° July, p. 1–5, 2011 (cf. p. 11).
- [50] R. RUICHONG, R. SNIEDER, L. GARGAB et A. SEIBI, “Modeling of seismic wave motion in high-rise buildings”, *Probabilistic Engineering Mechanics*, t. 26, n° 4, p. 520–527, 2011 (cf. p. 11).
- [51] D. M. SIRINGORINGO et Y. FUJINO, “System identification of suspension bridge from ambient vibration response”, *Engineering Structures*, t. 30, p. 462–477, 2008 (cf. p. 11).
- [52] A. MIKAEL, “Evaluation des paramètres physiques des bâtiments : amortissement, fréquences et modes de comportement des structures de génie civil : approche expérimentale”, thèse de doct., 2011 (cf. p. 11).
- [53] Y. YANG, C. W. LIN et J. D. YAU, “Extracting bridge frequencies from the dynamic response of a passing vehicle”, *Journal of Sound and Vibration*, t. 272, p. 471–493, 2004 (cf. p. 11).
- [54] G. a. PRIETO, J. F. LAWRENCE, a. I. CHUNG et M. D. KOHLER, “Impulse Response of Civil Structures from Ambient Noise Analysis”, *Bulletin of the Seismological Society of America*, t. 100, n° 5A, p. 2322–2328, sept. 2010 (cf. p. 11).
- [55] H. A. COLE, “On-line failure detection and damping measurement of aerospace structures by random decrement signatures”, rap. tech., 1973 (cf. p. 11).
- [56] P. KOPFF, “Renaissance d’une technique de l’analyse modale d’avant les transformées rapides : le décréement aléatoire”, *Recherche - Traitement du signal*, t. 8, p. 345–354, 1974 (cf. p. 11).

## 1.3 Travaux présentés

Le manuscrit articule la majorité des résultats obtenus au cours de ce doctorat et publiés dans diverses communications sous la forme de 2 articles de journaux [1, 2], 4 articles de conférences (3 internationales) [3, 4, 5, 6] et d'un brevet [7] :

**Chapitre 2 :** La démarche d'estimation de paramètres à partir de sources d'opportunité est abordée en relachant l'hypothèse de blancheur des sources. Des études théoriques et expérimentales conduisent à la construction de modèles d'observation qui tiennent compte des propriétés spectrales (en temps et en espace) des sources et à l'évaluation des performances d'estimation de différents paramètres physiques et géométriques ;

**Chapitre 3 :** Le protocole d'estimation de paramètres à partir de sources d'opportunité est évalué dans un contexte de Structural Health Monitoring. Des considérations théoriques et techniques liées à la mise en oeuvre d'un réseau de capteurs sont discutées ;

**Chapitre 4 :** De la physique, un modèle polynomial par morceaux est identifié pour la corrélation de Green. Une approche d'estimation bayésienne est mise en oeuvre sur ce modèle, permettant de court-circuiter l'emploi de l'identité de Ward pour estimer des paramètres physique et géométrique.

---

## Communications

- [1] M. CARMONA, O. MICHEL, J.-L. LACOUME, B. NICOLAS et R. VINCENT, “On Green’s correlation of Stokes’ equation”, *Wave Motion*, t. 56, p. 183–198, 2015 (cf. p. 19).
- [2] R. VINCENT, M. CARMONA, O. MICHEL et J.-L. LACOUME, “Inter-sensor propagation delay estimation using sources of opportunity”, *Signal Processing*, t. 118, p. 248–258, 2016 (cf. p. 19).
- [3] —, “A lower bound for passive sensor-network auto-localization”, dans *European Signal Processing Conference*, Lisbonne, 2014, p. 1965–1969 (cf. p. 19).
- [4] R. VINCENT, C. MIKAEL, O. MICHEL et J.-L. LACOUME, “Passive Acoustic Sensor Network Localization ; Application to Structure geometry Monitoring”, dans *7th European Workshop on Structural Health Monitoring*, Nantes, 2014, p. 197 –204 (cf. p. 19).
- [5] R. VINCENT, R. PERRIER, M. CARMONA et O. MICHEL, “Estimation bayésienne de la corrélation de Green ; application à l’estimation passive de distances inter-capteurs”, dans *GRETSI*, Lyon, France, 2015 (cf. p. 19).
- [6] R. VINCENT, M. CARMONA, O. MICHEL et J.-L. LACOUME, “SHM using acoustic sources of opportunity : monitoring Green’s correlations”, dans *International Workshop on Structural Health Monitoring*, 2015 (cf. p. 19).
- [7] —, “Brevet : Procédé de caractérisation d’un système disposé dans un milieu permettant la propagation d’une onde émise par un objet dudit système”, 2015 (cf. p. 19).



# Modèles, estimation et détection pour l'identification passive

---

Dans ce chapitre les résultats consistent majoritairement en la proposition de modèles d'observation pour le contexte passif d'estimation, issus de la physique et correspondant donc à présent au cadre de la propagation d'onde en milieu visco-acoustique. S'en suit une étude de la théorie de l'estimation et détection dans ce contexte. En particulier, nous insisterons sur l'influence du contenu spectral des champs mesurés, en temps et en espace, sur le comportement des estimateurs. Pour illustrer les propriétés avancées, un protocole expérimental est proposé, dans lequel nous expliquons comment exploiter le champ réverbéré de tout événement d'opportunité pour estimer trois quantités d'intérêt, à l'image de la démarche suivie avec l'oscillateur amorti. Nous nous intéresserons à l'estimation de la réponse impulsionnelle du milieu de propagation, à sa corrélation ainsi qu'à un paramètre géométrique, qui se trouve être la distance géodésique qui sépare deux capteurs. Les résultats expérimentaux sont bien entendu mis en regard des résultats théoriques issus de simulations.

## Sommaire

---

<b>2.1</b>	<b>Identification passive en acoustique aérienne . . . . .</b>	<b>23</b>
2.1.1	Equation de propagation et filtrage . . . . .	23
2.1.2	Corrélation de champs et corrélation de Green . . . . .	24
2.1.3	Identité de Ward en acoustique aérienne . . . . .	25
<b>2.2</b>	<b>Théorie de l'estimation/détection pour le contexte passif . . . . .</b>	<b>27</b>
2.2.1	Formulation du problème d'estimation . . . . .	27
2.2.2	Formulation du problème de détection . . . . .	31
<b>2.3</b>	<b>Expérimentation en acoustique confinée . . . . .</b>	<b>32</b>
2.3.1	Source d'opportunité, mais source exploitable . . . . .	32
2.3.2	Analyses des performances expérimentales . . . . .	39
<b>2.4</b>	<b>Caractérisation des performances dans un milieu contrôlé . . . . .</b>	<b>42</b>
2.4.1	Identité de Ward en milieu fini, conditions de Neuman . . . . .	42
2.4.2	Application : cas du parallélépipède rectangle . . . . .	46
<b>2.5</b>	<b>Discussion à propos de la distribution spatiale des sources . . . . .</b>	<b>52</b>
2.5.1	Modèle géométrique de sources d'opportunité . . . . .	52
2.5.2	Modèle d'observation pour la corrélation de Green . . . . .	53
2.5.3	Illustrations du filtrage spatial des sources . . . . .	55
2.5.4	Etude de performances conclusives . . . . .	58

---





## 2.1 Identification passive en acoustique aérienne

Dans cette première partie, le raisonnement suivi dans l'exemple de l'oscillateur amorti est repris et étendu à un milieu de propagation visco-acoustique, aux longueurs d'onde audibles par l'oreille humaine (de 15 mm à 15 m). L'objectif reste identique, il s'agit de trouver une relation qui permette de retrouver la réponse impulsionnelle du milieu de propagation à partir de la corrélation des champs mesurés, issus de sources d'opportunité. Dans un deuxième temps, il s'agit d'utiliser cette relation pour construire des estimateurs et d'en étudier les performances.

### 2.1.1 Equation de propagation et filtrage

Etudions donc le cas de la propagation acoustique, d'un champ scalaire donc, dans l'air. Les hypothèses faites sur le milieu sont celles de l'homogénéité et de l'isotropie, pour les longueurs d'ondes mesurées ; c'est-à-dire que les propriétés du milieu (température, pression, composition chimique) sont les mêmes en tout point de l'espace et dans toutes les directions de propagation. Le milieu est également linéaire, définissant ainsi qu'une excitation monochromatique résulte en un champ identiquement monochromatique. L'opérateur de propagation est par ailleurs une forme linéaire et il préserve donc les combinaisons linéaires.

L'air est un fluide visqueux, il résiste à sa mise en mouvement et donc à la propagation de la pression acoustique. L'opérateur de dissipation de type frottement fluide, illustré dans le cas de l'oscillateur amorti et amenant à l'identité (1.6), n'est pas adapté pour décrire le comportement de l'air. Un rappel complet sur la modélisation de la dissipation visqueuse dans l'air est proposé en Annexe B.1, inspirés des travaux [1]. Rappelons que l'équation de propagation considérée pour un fluide visqueux, voir [2] s'écrit :

$$\left[ \frac{\partial^2}{\partial t^2} - \alpha^2 \frac{\partial}{\partial t} \Delta - c_0^2 \Delta \right] p(t, \mathbf{x}) = f(t, \mathbf{x}). \quad (2.1)$$

Dans l'équation de propagation (2.1), c'est la quantité  $[\partial^2/\partial t^2 - \alpha^2 \partial/\partial t \Delta - c_0^2 \Delta]$  qui est appelée opérateur de propagation. Pour des champs de pression causaux, l'opérateur de propagation est inversible et il devient possible d'écrire le lien entre la source excitatrice du milieu et le champ résultant par une relation de filtrage généralisée en temps et en espace. C'est une relation qui, contrairement à l'équation de propagation, exhibe une transformation appliquée à la source  $f$  afin d'obtenir le champ résultant  $p$ . Cette transformation est le produit de convolution  $*$  en temps ( $t$ ) et en espace ( $\mathbf{x}$ ) :

$$p(t, \mathbf{x}) = \iint_{\mathbb{R}^3 \times \mathbb{R}} g(t, t', \mathbf{x}, \mathbf{x}') f(t', \mathbf{x}') dt' d\mathbf{x}' = [g * f](t, \mathbf{x}). \quad (2.2)$$

Dans l'expression (2.2), la fonction  $g(t, t', \mathbf{x}, \mathbf{y})$  est le noyau de l'opérateur de propagation et porte le nom de fonction de Green du milieu. Par analogie avec la théorie du filtrage, la fonction de Green est interprétée comme étant l'équivalent de la réponse impulsionnelle, généralisée aux dimensions de temps et d'espace. Elle se lit donc « réponse du milieu vue

en  $\mathbf{y}$  à l'instant  $t'$  sachant qu'une impulsion a été émise en  $\mathbf{x}$  à l'instant  $t$  ». Dans la suite, nous considérons le milieu invariant aux translations. L'invariance du milieu aux translations dans le temps est relative à la durée d'observation ; le milieu est dit invariant en temps si ses paramètres géométriques (son contour), chimiques (proportions des molécules) et physiques (température, pression) sont constants à l'échelle de la durée d'observation. Dans ce cas, la fonction de Green s'écrit en fonction du retard  $\Delta t = t - t' : g(\Delta t, \mathbf{x}, \mathbf{y})$ . L'invariance du milieu aux translations dans l'espace induit son caractère homogène et infini, tel que la fonction de Green ne dépend plus que de l'écart de position  $\Delta \mathbf{x} = \mathbf{x} - \mathbf{x}' : g(\Delta t, \Delta \mathbf{x})$ . La relation de filtrage se ré-exprime alors :

$$p(\Delta t, \Delta \mathbf{x}) = \iint_{\mathbb{R}^3 \times \mathbb{R}} g(\Delta t - t', \Delta \mathbf{x} - \mathbf{x}') f(t', \mathbf{x}') dt' d\mathbf{x}'. \quad (2.3)$$

Cette équation correspond à l'équation de filtrage identifiée pour l'oscillateur amorti, étendu aux domaines temporel et spatial.

### 2.1.2 Corrélation de champs et corrélation de Green

Comme pour le cheminement suivi avec l'oscillateur amorti, la prochaine étape consiste à relier la corrélation de la réponse du milieu à la réponse impulsionnelle. Travaillons donc sur la corrélation d'un champ. Faisons l'hypothèse que celui-ci est stationnaire, c'est-à-dire que ses statistiques temporelles et spatiales sont constantes, au moins à l'échelle de la mesure. Nous pouvons étendre la définition classique de la fonction de corrélation, pour une variable multivariée telle que le champ acoustique  $p(t, \mathbf{x})$ . Celle-ci, dans la mesure où le champ est assimilé à une variable aléatoire stationnaire et ergodique, peut être approximée par :

$$C_p(\Delta t, \Delta \mathbf{x}) \approx \lim_{T, \mathbf{X} \rightarrow \infty} \frac{1}{2T} \frac{1}{\text{vol}(\mathbf{X})} \int_{-T}^{+T} \int_{-\mathbf{X}}^{\mathbf{X}} p(t, \mathbf{x}) p(t + \Delta t, \mathbf{x} + \Delta \mathbf{x}) dt d\mathbf{x}. \quad (2.4)$$

Grâce à la relation de filtrage, Eq. (2.3), il devient possible de substituer le champ  $p(t, \mathbf{x})$  à sa relation avec la source  $f$  qui l'a généré et le milieu dans lequel le champ s'est propagé.  $C_p$  se ré-écrit en effet :

$$\iiint g(t', \mathbf{x}') C_f(\Delta t + t' - t'', \Delta \mathbf{x} + \mathbf{x}' - \mathbf{x}'') g(t'', \mathbf{x}'') dt' dt'' d\mathbf{x}' d\mathbf{x}''.$$

où en effet, en réunissant les termes concernés par l'intégration en  $t$  et  $\mathbf{x}$ , la corrélation de la source, aléatoire, notée  $C_f$ , est naturellement identifiée.

La corrélation du champ s'exprime donc en fonction de la corrélation de la source qui l'a généré et c'était un résultat attendu, car habituellement rencontré pour des filtres via la formule des interférences. Nous observons ici son extension pour un système régi par une équation de propagation spatio-temporelle, dans le domaine  $(t, \mathbf{x})$ . Des relations de filtrage relient les quantités en jeu et en posant  $t' = -t'$  et  $\mathbf{x}' = -\mathbf{x}'$ , la corrélation des champs prend la forme suivante :

$$C_p(\Delta t, \Delta \mathbf{x}) = \iint_{\mathbb{R} \times \mathbb{R}^3} [g_- * C_f] (\Delta t - t'', \Delta \mathbf{x} - \mathbf{x}'') g(t'', \mathbf{x}'') dt'' d\mathbf{x}'' \quad (2.5)$$

où  $g_-(t, \mathbf{x}) = g(-t, -\mathbf{x}) = g(-t, \mathbf{x})$ , est la version retournée dans le temps de la fonction de Green. Une dernière relation de filtrage est finalement identifiée, correspondant à l'extension de la formule des interférences pour un milieu de propagation, exprimée dans le domaine temporel et spatial :

$$C_p(\Delta t, \Delta \mathbf{x}) = [C_f * C_g](\Delta t, \Delta \mathbf{x}). \quad (2.6)$$

Dans l'Eq (2.6), l'expression  $C_g = g * g_-$  est une approximation de l'auto-corrélation de la fonction de Green. Cette quantité sera appelée **corrélacion de Green**, appellation proposée par J.-L. Lacoume [3] et que nous conservons par la suite. Elle sera notée  $C_g$ . Nous observons ici la formule des interférences généralisée pour des filtres variés en temps et en espace. Pour aller plus loin, notons à nouveau que si la source est blanche (stationnaire et d'énergie constante à toutes les fréquences et nombres d'onde), alors sa fonction d'auto-corrélation s'écrit

$$C_f(\Delta t + t' - t'', \Delta \mathbf{x} + \mathbf{x}' - \mathbf{x}'') = \delta(\Delta t + t' - t'')\delta(\Delta \mathbf{x} + \mathbf{x}' - \mathbf{x}'') \quad (2.7)$$

où  $\delta$  est la distribution de Dirac, élément neutre du produit de convolution, auquel cas la corrélation des champs (2.5) est la corrélation de Green

$$C_p(\Delta t, \Delta \mathbf{x}) = C_g(\Delta t, \Delta \mathbf{x}), \quad (2.8)$$

ce qui correspond à nouveau au pendant de l'oscillateur amorti pour un milieu de propagation. Observons qu'il ne reste plus que l'expression d'un filtrage entre la fonction de Green et sa version retournée dans le temps. Il n'y a plus de trace de la source qui a généré le champ ; elle a été « oubliée ». La corrélation de Green ne contient d'information que sur le milieu de propagation.

### 2.1.3 Identité de Ward en acoustique aérienne

A ce stade, nous avons identifié ce qu'est la corrélation du champ stationnaire et son lien avec la corrélation de Green. Il reste alors à trouver la relation, pour un milieu de propagation visco-acoustique, qui permettra d'identifier la fonction de Green à partir de la corrélation de Green. Dans l'état de l'art, voir par exemple [4, 5, 6], une identité a été proposée pour expliciter un lien fonctionnel entre cette corrélation de Green et la fonction de Green d'un milieu à dissipation constante. Dans ses travaux, Mikael Carmona [7] traite le cas visco-acoustique et nous en livrons, à la fin de cette section, une version qui tient compte des propriétés spectrales du champ mesuré.

L'identité de Ward visco-acoustique est la suivante :

$$\frac{\partial^3}{\partial t^3} C_g(\Delta t, \mathbf{x} - \mathbf{y}) \approx \frac{c_0^2}{\alpha^2} \text{Odd}[g](\Delta t, \mathbf{x} - \mathbf{y}). \quad (2.9)$$

où la notation  $\text{Odd}[g] = \frac{1}{2}(g - g_-)$  exprime la partie impaire de la fonction de Green.

Cette identité, pendant de l'Eq. (1.6) pour l'oscillateur amorti, est construite en s'appuyant sur les hypothèses suivantes. Le système est invariant aux translations en temps et en espace

et la source est stationnaire, blanche en temps et en espace. Par ailleurs, la dissipation est modélisée par l'opérateur  $-\alpha^2 \partial / \partial t \Delta$  et est faible, ce qui intervient pour obtenir un résultat approché (opérateur  $\approx$ ), mais facilement exploitable. Une preuve de cette identité est proposée en Annexe B.2. Elle s'appuie sur son calcul préliminaire dans le domaine fréquence-espace  $(\omega, \mathbf{x})$  puis son inversion par l'application du théorème des résidus pour la retrouver dans le domaine temps-espace  $(t, \mathbf{x})$ .

Pour conclure, notons que la combinaison de la formule des interférences (2.8) et de l'identité de Ward visco-acoustique (2.9) permet de généraliser la reconstruction de la fonction de Green du milieu en fonction des propriétés spectrales des champs mesurés :

$$\frac{\partial^3}{\partial t^3} C_p(t, \mathbf{x} - \mathbf{y}) \times u(t) \approx \frac{c_0^2}{\alpha^2} [C_f * g](t, \mathbf{x} - \mathbf{y}) \times u(t). \quad (2.10)$$

Notons premièrement que c'est la partie impaire de la fonction de Green qui est retrouvée par l'identité (2.9), ce qui ne constitue pas un obstacle pour retrouver la fonction de Green complète, puisque cette dernière est causale. Nous illustrons cette notion par la Fig. 2.1.

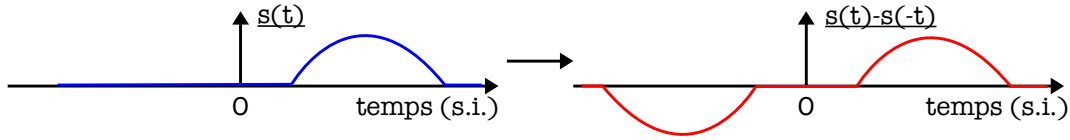


FIGURE 2.1 – Partie impaire d'une fonction causale

Deuxièmement, la dissipation acoustique est faible dans la plage des fréquences audibles. En ce sens en effet, un parcours de 1000 mètres conduit à une atténuation géométrique (conservation énergétique de l'énergie du front d'onde) de 60 dB et de 2 dB (à 1kHz) pour la dissipation due à la viscosité de l'air. Ainsi, l'hypothèse de faible viscosité est justifiée et se traduit par la linéarisation du nombre d'onde  $k$  :

$$k = \frac{k_0}{\sqrt{1 + jk_0\alpha^2}} \approx k_0, \text{ avec } k_0 = \frac{\omega}{c} \quad (2.11)$$

où l'on retrouve que  $\alpha^2 \ll c/\omega$ , signifiant que la dissipation visqueuse est très faible à l'échelle de la distance parcourue par l'onde pendant 1 sec. Si la viscosité est faible, notons que sans elle, l'identité de Ward Eq. (2.9) n'existe tout simplement pas.

Finalement, si le champ source est blanc en temps et en espace, la corrélation de Green peut théoriquement être reconstruite avec un résidu arbitrairement petit, ne dépendant que de la longueur des signaux considérés et du bruit de mesure à l'acquisition. Puisque la relation de Ward est une identité, la fonction de Green peut être estimée de la même manière. Si en revanche la source n'est pas blanche en temps, alors la fonction de Green reconstruite dépend du contenu spectral de la source : il n'est possible de reconstruire la fonction de Green que sur la bande où de l'information est présente. Deuxièmement, c'est la corrélation de la source qui intervient pour filtrer la fonction de Green théorique : cette corrélation est caractérisée par exemple par une longueur caractéristique (décroissement de -60dB par exemple) qui sera d'autant plus grande que la source ne sera pas blanche sur la bande considérée.

## 2.2 Théorie de l'estimation/détection pour le contexte passif

Dans le contexte passif d'estimation, les sources sonores ne sont pas maîtrisées par l'utilisateur. Afin de reconstruire malgré tout des paramètres qui caractérisent le milieu de propagation, il faudra leur formuler des estimateurs.

Nous nous intéresserons à l'estimateur d'un paramètre géométrique. En sismologie, connaissant la distance qui sépare deux stations, estimer la célérité de propagation de l'onde à partir du temps que l'onde a mis pour se propager d'une station à l'autre permet d'imager l'intérieur de la Terre et de localiser la profondeur et la composition des couches les plus profondes de la planète. Dans notre cas, la célérité d'une onde acoustique aérienne est connue et nous chercherons plutôt à estimer la distance qui sépare les capteurs. Etendue à l'échelle d'un réseau de capteurs, cette information ouvre la porte à la thématique de recherche de la « reconstruction de géométries ».

### 2.2.1 Formulation du problème d'estimation

#### 2.2.1.1 Modèle d'observation pour la fonction de Green

Si les dimensions du milieu de propagation sont bien plus grandes que la longueur d'onde considérée, alors modéliser la propagation d'un front d'onde par celle d'un ensemble de rayons obéissant aux lois de Snell-Descartes est une approximation raisonnable, voir [8] par exemple. Ce régime de propagation s'appelle le régime « hautes fréquences » et il est couramment employé, en acoustique aérienne, pour modéliser le comportement acoustique d'une salle. Dans ce régime de propagation, une réponse impulsionnelle est conventionnellement modélisée par un ensemble de retards  $t_i$  correspondant aux temps requis par l'onde pour se propager le long des trajets géométriques reliant la source au capteur. Ces retards sont associés à une amplitude  $a_i$  de l'onde décroissante avec la distance, due à l'atténuation géométrique, par dissipation, par les éventuelles réflexions (pour un milieu borné), couplages et autres phénomènes diffuseurs d'énergie. Illustrée Fig. 2.2, la fonction de Green peut être modélisée par :

$$g(t, \mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\infty} a_i \delta(t - t_i), \text{ avec } |a_i| \leq |a_0|. \quad (2.12)$$

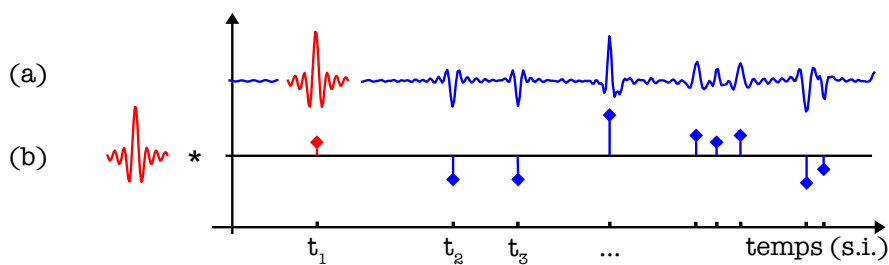


FIGURE 2.2 – Modélisation de la fonction de Green en régime de propagation haute fréquence

En milieu de propagation infini homogène, ce modèle se résume au premier temps de propagation  $t_1$  tel que  $g(t, \mathbf{x} - \mathbf{y}) = a_1 \delta(t - t_1)$  et peut, dans une certaine mesure, être appliqué en milieu fini. C'est en effet le cas si par exemple l'horizon temporel d'observation est court ( $t \leq t_2$ ). Cette simple considération permet de conserver le modèle d'observation pour la corrélation de Green et la fonction de Green élaboré en milieu infini, dans le contexte réaliste et pratiquement rencontré du milieu fini. Ainsi, à partir de l'Eq. (2.10), un modèle d'observation pour la fonction de Green dans un contexte passif peut être formulé par :

$$g(t, \mathbf{x}, \mathbf{y}) = \frac{c_0^2}{\alpha^2} C_f(t) * \delta(t - t_1(|\mathbf{x} - \mathbf{y}|)) + w(t), \quad (2.13)$$

où  $w(t)$  est un bruit de modèle.

Que le contexte d'estimation soit actif ou passif, un estimateur  $\hat{t}_1$  pour  $t_1$  peut être l'amplitude maximale de la fonction de Green, soit ici :

$$\hat{t}_1 = \arg \max_t y(t). \quad (2.14)$$

Ainsi, il apparait maintenant comment la corrélation de la source  $C_f$  est au coeur du problème d'estimation dans un contexte passif. Cette dernière sert (1) à estimer la corrélation de Green, (2) puis la fonction de Green par application de l'identité de Ward et enfin (3) tout autre paramètre qui pourrait en découler, comme  $t_1$  par exemple. Ce qu'il faut retenir, c'est que tout estimateur dans un contexte passif d'estimation voit ses performances conditionnées par la capacité de l'utilisateur à estimer correctement la corrélation de Green à partir des champs mesurés. Ses performances ne seront vraisemblablement que la conséquence de la propagation de l'erreur commise à l'estimation de la corrélation de Green.

### 2.2.1.2 Bruit utile, bruit de mesure et résidu d'estimation

Faisons à présent un point pour préciser ce qu'est une information utile pour l'identification d'un milieu de propagation visco-acoustique dans un contexte passif d'estimation. Premièrement, notons que pour appliquer l'identité de Ward Eq. (2.9), le champ source a été modélisé par une variable aléatoire de densité spectrale de puissance spatiale constante. Le champ est par ailleurs supposé stationnaire, caractérisé par une fonction d'auto-corrélation en temps donnée, ce qui permet une formalisation assez générique du problème. Comme expliqué dans la section précédente 2.2.1.1, les différents modèles d'observation pour le contexte passif d'estimation que sont la corrélation de Green, la fonction de Green et le temps de propagation, s'expriment tous comme une cascade en série de transformations appliquées aux champs mesurés par les capteurs. Les estimateurs qui sont construits à partir de ces modèles d'observations propagent donc les incertitudes issues des étapes précédentes.

La toute première incertitude est introduite à l'étape de la mesure, où la chaîne d'acquisition apporte un bruit électronique qui vient s'ajouter au bruit du capteur (qui dépend des conditions d'utilisation et de la qualité du capteur) et le passage dans le monde numérique

introduit à son tour des erreurs de quantification et d'échantillonnage :

$$\begin{cases} x[n] = p(t_0 + nT_s, \mathbf{x}) + w_x[n] \\ y[n] = p(t_0 + nT_s, \mathbf{y}) + w_y[n] \end{cases} \quad (2.15)$$

Traditionnellement, le bruit total, noté ici  $w_x[n]$  ou  $w_y[n]$ , résultant sur l'observation échantillonnée et quantifiée de la grandeur physique est modélisé par une séquence blanche de bruit additif gaussien centré. C'est un modèle souvent réaliste qui permet l'écriture analytique de certains concepts importants comme les estimateurs de maximum de vraisemblance ou diverses bornes sur les performances des estimateurs. Ensuite intervient la corrélation numérique pour estimer la corrélation du champ, puis l'application de l'identité de Ward pour obtenir la fonction de Green. Finalement intervient la recherche du maximum d'amplitude.

### 2.2.1.3 Corrélateur

Un estimateur pour la corrélation, applicable au modèle d'observation Eq. (2.13), peut être construit en calculant l'espérance temporelle et spatiale des champs à l'aide d'une fenêtre glissante d'analyse. Ainsi, la corrélation estimée est la moyenne d'ensemble des corrélations estimées sur chacune des  $K$  fenêtres d'analyse :

$$\hat{C}_p[n] = \frac{1}{K} \sum_{k=1}^K \frac{1}{(b_k - a_k)\sqrt{E_k}} \sum_{l=a_k}^{b_k-n} x[l]y[l-n]. \quad (2.16)$$

où les  $a_k$  et  $b_k$  définissent les extrémités de la fenêtre glissante d'analyse. Cet estimateur classique est biaisé mais à une variance faible, voir [9]. Le biais, convolutif (fenêtre de Barlett), est compensable par ailleurs par l'utilisation d'une fenêtre d'apodisation (par exemple de Blackman). La constante de normalisation  $E_k = \sqrt{\sum_k x[k]^2 \sum_k y[k]^2}$  fait apparaître l'énergie des champs dans la  $K$ ième fenêtre d'analyse. Cette normalisation contre la variabilité de l'énergie des champs mesurés en pratique, étant donné que les sources qui les génèrent ne sont pas contrôlées. En procédant ainsi, leurs contributions sont égalisées mais l'information d'amplitude est perdue. C'est un choix qui peut ne pas être acceptable pour estimer certains paramètres. Dans notre cas et pour les paramètres que nous visons à reconstruire, nous verrons que c'est un choix justifié.

### 2.2.1.4 Application de l'identité de Ward

La relation qui relie la corrélation des champs à la fonction de Green du milieu de propagation, entre les points de mesure, est une relation de filtrage. Ici, nous l'avons approchée par une relation de filtrage temporel, et correspond à une dérivation. La dérivation peut numériquement être implémentée avec des soustractions et c'est le choix que nous avons fait ici, de telle sorte que :

$$\hat{g}[n] = \hat{C}_p[n] - 3\hat{C}_p[n-1] + 3\hat{C}_p[n-2] - \hat{C}_p[n-3]. \quad (2.17)$$



Dans l'Annexe B.3, il est montré qu'en dérivant trois fois une séquence blanche de bruit gaussien centré, en utilisant cette technique de filtrage basique par soustraction, le résultat obtenu est une séquence également blanche gaussienne, mais de variance  $\binom{6}{3} = 20$  fois plus grande que la séquence originelle<sup>1</sup>. En revanche, cette étape est nécessaire pour faire émerger la fonction de Green. La variance du bruit de modèle augmente, c'est certain, mais la quantité d'intérêt, c'est-à-dire la fonction de Green, émerge par cette transformation, si bien que le RSB lui, n'évolue pas pour autant de manière négative.

### 2.2.1.5 Estimateur de l'argmaximum

La fonction d'auto-corrélation de la source est symétrique, et son maximum d'amplitude se trouve au retard nul. Le maximum d'amplitude de la fonction est donc localisé en  $t_1$  et l'estimateur du temps de propagation  $\hat{t}_1$  peut donc être la position de l'amplitude maximum de la fonction de Green estimée. Etant donné le modèle d'observation 2.13, classique, nous savons que sa variance dépend de la largeur de bande de la fonction reconstruite. Pour illustrer ce propos, considérons le cas d'un champ source blanc sur une bande  $B$ . La fonction de Green reconstruite est donc modélisée par :

$$\hat{g}(t) = A \operatorname{sinc}(B(t - t_1)) + w(t). \quad (2.18)$$

Ce modèle, classique, est également générique dans le sens où nous avons montré qu'il permet de représenter autant le cas passif d'estimation que le contexte actif, plus habituel. C'est la force de l'approche, où l'effort est porté sur l'estimation de la fonction de Green. De cette manière, les outils classiques de traitement du signal peuvent être appliqués au contexte passif d'estimation, à ceci près que les RSB ne sont pas les mêmes.

Dans l'Eq. (2.18), le résidu d'estimation  $w(t)$  est modélisé gaussien, de densité spectrale de puissance constante égale à  $N_w$  sur  $B$  et nulle ailleurs. Notons  $1/(2B_s)$  la période d'échantillonnage. Pour ce modèle de bruit additif gaussien, une borne classique sur la variance de l'estimateur du temps de propagation est la borne de Cramér-Rao. La borne de Cramér-Rao pour  $t_1$ , présentée par exemple dans [10], dérivée de l'inégalité de Cauchy-Schwarz, s'écrit :

$$\mathbb{E}[(\hat{t}_1 - t_1)^2] \geq \frac{N_w}{\frac{4}{3}\pi^2 A^2 B_s B^2}. \quad (2.19)$$

Dans la borne Eq. (2.19), la quantité  $\frac{A^2 B}{4N_w}$  est assimilée à un RSB (voir [9] pour une justification approfondie) et  $\frac{16}{3}\pi^2 B B_s$  correspond au carré la largeur de bande moyenne du signal, telle qu'elle est définie dans [11] par exemple. De manière classique, cette borne inférieure théorique indique que la confiance sur l'estimation augmente avec la largeur de bande moyenne du signal et avec le RSB.

---

1.  $\binom{a}{b}$  est le coefficient binomial, «  $b$  parmi  $a$  »

### 2.2.2 Formulation du problème de détection

Après la description du problème d'estimation, complétons l'étude en précisant le problème de détection auquel nous faisons face. Concrètement, il s'agit de localiser le temps de propagation  $t_1$ , que nous espérons se trouver au maximum d'amplitude de  $\hat{g}[n]$ . La probabilité que cet événement se produise est une probabilité de détection standard en télécommunication, voir [11] pour une mise en contexte. Elle est illustrée Fig. 2.3 pour rappel. Pour décrire ce problème, nous avons besoin de calculer la probabilité que l'amplitude d'un bruit gaussien centré de variance  $\sigma^2$  soit plus grande qu'un seuil  $A$ , ce qui s'écrit :

$$\mathbb{P}[w_i \geq A] = \frac{1}{2} \left( 1 - \operatorname{erf} \left( \frac{\sqrt{2}}{2} \frac{A}{\sigma_i} \right) \right), \quad (2.20)$$

où  $\operatorname{erf}(x) = 2/\sqrt{\pi} \int_0^x \exp(-t^2) dt$  est la fonction erreur.

Cette probabilité  $\mathbb{P}[w_i \geq A]$  correspond à la probabilité d'erreur de détection à l'échelle d'un échantillon, qui augmente comme prévu lorsque le RSB  $A^2/\sigma^2$  diminue. La preuve de ce résultat est proposée en Annexe B.4.

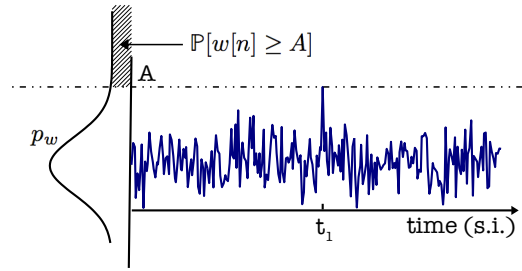


FIGURE 2.3 – Modèle de la fonction de Green  $\hat{g}(t) = A\delta(t - t_1) + w(t)$  estimée et probabilité d'erreur d'extraction du temps de propagation par argmax.

En conclusion, en décortiquant le problème d'estimation dans un contexte passif, il est apparu que le modèle d'observation pour la fonction de Green ou encore le temps de propagation  $t_1$  sont les mêmes que dans le contexte actif d'estimation. La véritable différence se trouve dans le rapport signal à bruit, où le contexte passif d'estimation accuse une cascade de transformations propageant les incertitudes de la mesure des champs. Pour rappel, les transformations en jeu sont : un estimateur de corrélation normalisé, un estimateur de dérivation à l'ordre 3 et la localisation de l'argmax. Les problèmes de détection/estimation décrits sont les éléments de base sur lesquels construire des bornes classiques telles que la borne de Cramér-Rao présentée Eq. (2.19), qui s'appuie sur l'inégalité de Cauchy-Schwarz, les bornes de type Ziv-Zakai, voir [11] (inégalités de Chebychev pour les moments d'ordre 2 ou de Markov pour les moments supérieurs).

Dans la suite, la borne de Cramér-Rao présentée Eq. (2.19) et la probabilité d'erreur Eq. (2.20) seront comparées à des résultats expérimentaux et de simulation. Plusieurs expériences seront menées. La première permettra d'introduire le protocole d'estimation suivi et d'illustrer les performances d'estimation en regard de la largeur de bande temporelle des

signaux. La deuxième expérience aura lieu dans un milieu où la comparaison entre les fonctionnelles de Green expérimentalement reconstruites et leur pendant théorique sera possible. Cette comparaison permettra de statuer sur notre niveau de modélisation et de réaliser l'étude de la convergence des estimateurs.

## 2.3 Expérimentation en acoustique confinée

Dans cette section, les résultats théoriques présentés jusqu'à maintenant sont illustrés par une expérience. Les trois quantités que nous chercherons à estimer seront la corrélation de Green, la fonction de Green puis le temps que met une onde acoustique pour se propager entre deux microphones. Le protocole d'estimation est présenté, accompagné de l'étude de ses performances. En particulier, nous nous consacrerons exclusivement à la caractérisation des performances d'estimation du temps de propagation  $t_1$ , en fonction de la largeur de bande des signaux mesurés. A ce stade, nous ne possédons pas de modèle pour la corrélation de Green.

### 2.3.1 Source d'opportunité, mais source exploitable

Pour mener les calculs et obtenir l'identité de Ward Eq. (2.9), l'hypothèse de blancheur spatiale et de stationnarité du champ source était nécessaire. En pratique, le bruit sismique utilisé en géophysique, c'est-à-dire la coda sismique, est issu de micro-séismes, des mouvements des plaques tectoniques, etc. En acoustique audible, nous chercherons à exploiter la partie réverbérée d'événements acoustiques non contrôlés. En pratique, les événements non contrôlés que nous manipulerons le plus seront des impulsions, utilisées pour deux raisons. Premièrement parce qu'elles sont couramment rencontrées dans la vie de tous les jours, par exemple sous la forme du choc de deux objets, de bruits de pas, de portes qui claquent, de clefs sur un bureau. Deuxièmement, parce que leur contenu spectral est blanc sur une bande donnée, permettant d'obtenir une réponse de la part du milieu sur cette bande et donc d'être informative (à l'opposé d'une onde monochromatique). Ainsi, d'après les suggestions de Aki [12] et Campillo [13], les codas acoustiques de réponses impulsives seront les signaux utilisés pour émuler la réponse du milieu à une source blanche en espace.

#### 2.3.1.1 La coda, ce que l'on sait

Je propose dans cette partie, de manière rapide, un récapitulatif des propriétés connues principales de la coda acoustique. Commençons par statuer sur le fait que la notion de blancheur spatiale est, à l'heure actuelle, non caractérisée. La communauté scientifique a bel et bien porté un vif intérêt à cette discussion, pendant une large décennie, mais n'a pu apporter que des éléments de réponses partiels. Concernant les propriétés que le champ diffus semble posséder pour être exploitable dans un cadre passif d'estimation, notons que Weaver [14] utilise une base modale pour décomposer les signaux et appelle le champs diffus : « bruit ambiant », qu'il modélise en répartissant de manière équitable l'énergie du signal dans tous les modes.

Cette notion d'équipartition trouve son équivalent avec les ondes sismiques, où le ratio des énergies des ondes P et S est stabilisé ; comportement induit par les multiples interactions avec les hétérogénéités du milieu [15]. De son côté, Snieder [16] se concentre davantage sur la notion d'isotropie des flux d'énergie pour quantifier un état de mélange. Des analyses proposées par Tromp, Hanasoge [17], rendent compte de l'impact de la distribution des sources dans le milieu sur les résultats de reconstruction des paramètres d'intérêt. Il est clair que c'est un paramètre qui, s'il est mis de côté, rendra incomplète n'importe quelle analyse quantitative. Wapenaar [18] quant à lui utilise des théorèmes de réciprocité, qu'il relie à la notion de renversement temporel et que l'on doit principalement aux travaux de M. Fink [19]. Parmi les débats animant la communauté, soulignons celui de la nécessité de la dissipation pour que l'identification passive soit possible. Notons également que le « désordre » a longtemps été considéré exclusivement bénéfique à l'identification passive, ayant la capacité de « blanchir » le champ. Des résultats récents montrent au contraire qu'un excès de diffusion peut être nuisible, justifié par le fait qu'une perte excessive d'énergie (par diffusion) peut empêcher l'onde d'explorer suffisamment le milieu pour devenir informative sur celui-ci.

Ceci étant dit, de nombreux enseignements ont été tirés des résultats expérimentaux et les champs amenant à des identifications réussies ont été caractérisés dans une certaine mesure.

Le milieu étant linéaire supposé invariant à l'échelle des mesures, une source stationnaire entraîne une réponse stationnaire (en écartant les régions transitoires dues à la dynamique du filtrage). En pratique, c'est une situation rencontrée en géophysique avec le « murmure » de la Terre, la coda sismique, c'est-à-dire l'activité sismique excluant les événements sismiques majeurs. En acoustique, la partie d'une réponse impulsionnelle acoustique qui est dans un état de réverbération avancée sera appelée coda acoustique. Qualifiée par certains d'incohérente [20], elle véhicule la notion de perte de mémoire de source, voir par exemple [21]. La coda n'est pas un champ stationnaire puisqu'elle est constituée d'une onde qui se propage, subissant une atténuation d'amplitude géométrique (conservation d'énergie du front d'onde) et propagative (dissipation). Elle peut cependant être considérée comme telle sur quelques centièmes de secondes, ce qui offre déjà un horizon de propagation de plusieurs mètres. Associée aux phénomènes de dissipation et d'atténuation, il faut noter par ailleurs que la bande passante du champ réverbéré, d'abord à l'image de la source qui l'a généré, réduira au fur et à mesure de la propagation.

En acoustique aérienne, une expertise avérée de la communauté concerne le traitement acoustique des salles. Notons en particulier que le champ réverbéré est une composante physique très étudiée. La réverbération, construite principalement par les quatre phénomènes que sont la diffusion, la réflexion, la diffraction et la réfraction, est en pratique réduite au strict minimum pour assurer une écoute confortable. À défaut de proposer un cadre d'exploitation pour notre cadre d'estimation passive, la taille écrasante de la littérature dans le domaine indique que la modalité acoustique aérienne est bien en mesure de fournir un champ diffus, potentiellement exploitable dans un cadre d'estimation passive.

### 2.3.1.2 Modélisation d'une coda

La coda est un champ propagé. Dans un régime de propagation haute fréquence (lorsque la longueur d'onde est plus petite que les dimensions caractéristiques du milieu), le parcours d'un front d'onde est bien représenté par le parcours de rayons [8]. La réponse impulsionnelle est alors constituée d'un ensemble de retards associés à des amplitudes dont l'enveloppe est dictée par l'atténuation géométrique et dissipative :

$$coda(t, r) \propto \frac{E_0}{r} \exp(-\mathcal{D}(\omega)r) \Pi(r) \quad (2.21)$$

où  $E_0$  est l'énergie du champ source,  $\mathcal{D}$  est l'opérateur de dissipation ( $\propto \omega^2$  dans l'air par exemple).  $\Pi$  est une variable modélisant les trajets des rayons. Une réalisation résulte en un peigne de Dirac, dont la densité dépend de  $r$ , la distance parcourue. Elle est modélisée par une loi  $p(r)$  qui dépend des bords du milieu et de la géométrie source-capteur. En acoustique des salles, une forme générale de  $p(r)$  est donnée, tenant compte du volume de la salle :

$$p(r) = 4\pi c_0 \frac{r^2}{V}. \quad (2.22)$$

où  $c_0$  est la célérité de l'onde.

Pour donner un ordre de grandeur, considérons une salle de 250 m<sup>3</sup> et un état où le front d'onde a déjà parcouru 100 mètres (soit en moyenne 15 rebonds). Lorsque le front d'onde se propagera 10 cm plus loin, chaque volume infinitésimal du milieu sera visité en moyenne 36000 par le front d'onde. Après s'être propagé sur une telle distance, l'onde vivra en moyenne encore deux fois plus longtemps avant de ne plus être distinguable du bruit d'une chaîne d'acquisition acoustique standard. Ainsi, il est clair que très rapidement dans le champ réverbéré, la densité d'interférences des fronts d'ondes est tellement élevée qu'en invoquant le théorème centrale-limite, il est naturel de justifier que spatialement, le champ observé soit un processus aléatoire continu Gaussien et centré. Des approches dans l'état de l'art ont été développées pour quantifier à quel point un signal est Gaussien. Ainsi, des estimations de moments d'ordre supérieur comme le Skewness et le Kurtosis sont parfois employées pour décrire l'état de réverbération du champ acoustique, voir par exemple [22].

Par ailleurs, la densité modale (nombre de modes compris dans 1 Hz) suit une loi :

$$m(f) = 4\pi V \frac{f^2}{c_0^3}, \quad (2.23)$$

et il s'avère qu'à cause de la dissipation, les modes possèdent une certaine largeur de bande. Ainsi, au dessus de la fréquence de Schroeder (quelques centaines de Hz dans une salle de type bureau), la couverture spectrale est complète et peut potentiellement accueillir de l'énergie sur toute la bande passante de la chaîne d'acquisition.

Ainsi, pour une source impulsionnelle, de l'énergie est fournie au champ sur une bande donnée (par exemple un claquement de main fourni typiquement de l'énergie jusqu'à 15 kHz), et d'après tout ce qui vient d'être dit, le champ réverbéré obéit donc à toutes les hypothèses

pour être modélisé comme un champ gaussien centré, localement stationnaire et même blanc dans une bande dictée par la dissipation et l'avancement de la propagation. La Fig. 2.4 présente le spectrogramme d'une réponse impulsionnelle acoustique dans une salle de type bureau, illustrant la perte d'énergie et la place occupée par la coda dans la réponse.

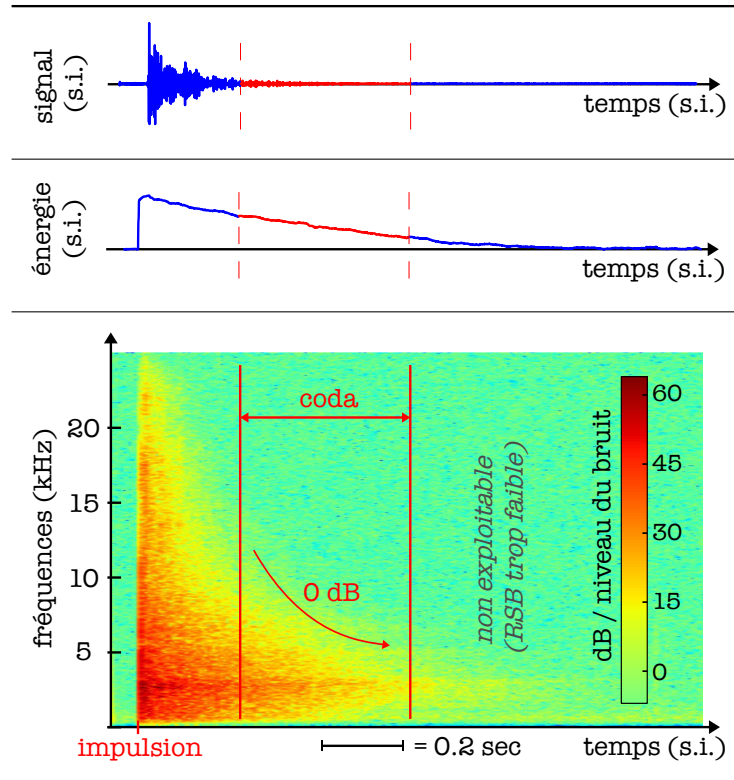
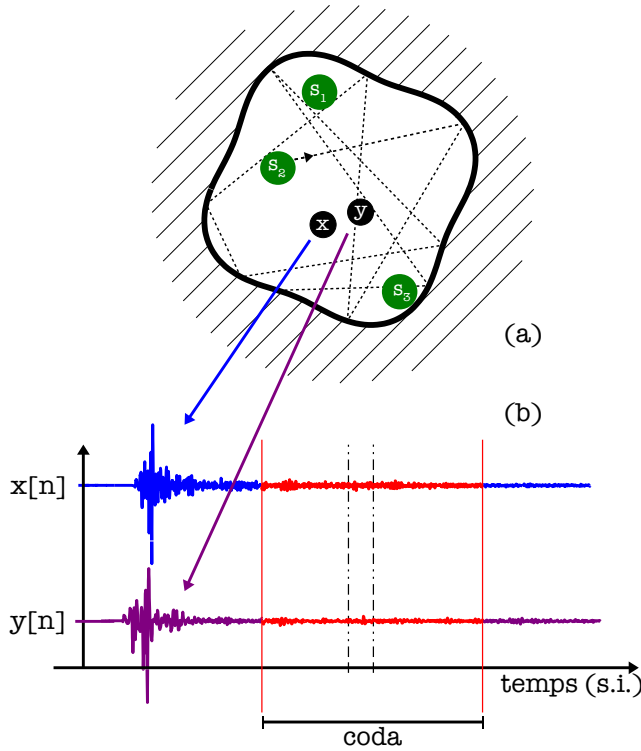


FIGURE 2.4 – Localisation d'une coda dans une réponse impulsionnelle aérienne.

Pour clôturer cette partie, je propose une ouverture à la notion de blancheur spatiale. L'équation de Helmholtz est l'analogue formelle stricte de l'équation de Schrödinger pour des ondes acoustiques, la géométrie du milieu étant équivalente aux conditions classiques aux limites du mouvement d'une particule. Dans leurs travaux, voir [23], Legrand et Mortessagne étudient la propagation de rayons dans un billard chaotique, paradigme géométrique à une cavité dans laquelle une onde acoustique se propage, pour caractériser certaines propriétés statistiques ; par exemple les patterns modaux et la densité spectrale du champ. La formule des traces de Gutzwiller, construite à partir des orbites périodiques dans les billards chaotique est l'outil utilisé pour faire ce lien. Ainsi, des notions telles que l'ergodicité spatiale des modes et la rigidité spectrale sont illustrées, véhiculant une information sur la corrélation spatiale et spectrale du champ. Selon moi, la notion de blancheur spatiale peut être étudiée au travers de l'outil connu sous le nom de facteur de forme semi-classique, présenté dans l'article. Cette étude permettrait de définir des arguments sur lesquels se baser pour choisir un temps de mélange, un début à la coda acoustique. C'est une des perspectives de mes travaux.

## 2.3.1.3 Description du protocole

Dans l'expérience proposée, des microphones sont positionnés dans une salle de type bureau meublé, de dimensions  $10\text{m} \times 6\text{m} \times 3.5\text{m}$  et de longueur caractéristique  $\sqrt[3]{10 \times 6 \times 3.5} = 8\text{ m}$ . Les microphones sont distants de 48 cm et le temps que met l'onde pour parcourir cette distance est 1.4 ms. Les capteurs enregistrent les séries temporelles  $\mathbf{x}[n]$  et  $\mathbf{y}[n]$  à la fréquence d'échantillonnage 44.1 kHz. Les sources sont des claquements de main, répartis aléatoirement dans l'espace et avec des puissances différentes. Les impulsions générées sont observées par les capteurs, voir Fig. 2.5, et les codas en sont extraites pour estimer les quantités de Green.

**Légende (Figure 2.5)**

La scène d'identification passive est composée de :

- Un milieu de propagation ;
- Des capteurs observant les champs se propageant dans le milieu ;
- Une ou plusieurs sources  $s_i$ , qui sont des sources d'opportunité ; elles sont non contrôlées par l'utilisateur, c'est-à-dire que leur position, spectre et instants de déclenchement ne sont pas sous le contrôle de l'utilisateur.

La réponse des capteurs, partie (b), produit des données que le protocole d'identification passive cherchera à exploiter pour estimer des paramètres relatifs au milieu de propagation ou au réseau de capteurs.

FIGURE 2.5 – Schéma de principe.

Les événements impulsionnels sont extraits en s'appuyant sur l'énergie instantanée du signal, décrite dans l'algorithme 1, et résumée par deux étapes :

- (1) Le début d'un événement de type impulsion est détecté lorsque l'énergie du signal s'élève brusquement (règle sur l'énergie du signal et sur sa variation) ;
- (2) Une fois le début de cet événement localisé, la fin de l'événement est décrétée lorsque le rapport signal à bruit passe en dessous d'un certain seuil, par exemple lorsque l'énergie du signal retourne au niveau du bruit de mesure.

Pour déterminer où débute la coda acoustique, aucune règle absolue ne fait foi dans l'état de l'art, pour la simple raison que la propriété de blancheur spatiale n'est pas maîtrisée, alors que son rôle est essentiel pour l'application de l'identité de Ward. Basiquement, retenons que c'est un état de réverbération avancé qui est recherché, pour se défaire des trajets primaires sporadiques fortement corrélés entre eux. Si rechercher une limite entre deux régimes de pro-



**Algorithme 1:** Détection d'évènement impulsif

---

- tronçonner le signal en fenêtres avec overlap;	1
- initialiser le booléen <i>isEvt</i> = <b>false</b> ;	2
- initialiser les seuils $\lambda_1, \lambda_2, \lambda_3$ ;	3
<b>forall the</b> <i>fenêtre</i> <b>do</b>	4
- calculer l'énergie $E_i$ du signal sur la fenêtre <i>i</i> ;	5
<b>if</b> $E_i \geq \lambda_1$ <b>and</b> <i>isEvt</i> == <b>false</b> <b>then</b>	6
<b>if</b> $E_i - E_{i-1} \geq \lambda_2 \geq 0$ <b>then</b>	7
- marquer l'instant courant comme le début d'un évènement;	8
- <i>isEvt</i> = <b>true</b> ;	9
<b>end</b>	10
<b>end</b>	11
<b>if</b> <i>isEvt</i> == <b>true</b> <b>then</b>	12
<b>if</b> $E_i \leq \lambda_3$ <b>then</b>	13
- marquer l'instant courant comme la fin d'un évènement;	14
- <i>isEvt</i> = <b>false</b> ;	15
<b>end</b>	16
<b>end</b>	17
<b>end</b>	18

---

pagation représente une finalité théorique attrayante, nous n'en sommes pas encore capable. L'objectif ici est plutôt de fournir un critère qui permette d'extraire un signal exploitable par application de l'identité de Ward, et non pas le meilleur segment, qui serait pas exemple le plus long possible. L'approche proposée ici « esquivé » donc la difficulté, reconnue par ailleurs dans la littérature : nous comparons plusieurs temps caractérisant la propagation de la pression acoustique dans une salle, construits sur des considérations différentes, et nous préférons en déduire une règle générale robuste. En acoustique, il a été montré que la densité de trajets  $D(t)$  arrivant sur un capteur, lorsque la source est une impulsion, augmente avec le carré du temps

$$D(t) = 4\pi c_0^3 \frac{t^2}{V}, \quad (2.24)$$

et nous présentons trois temps représentatifs d'une limite de résolution définis à partir de cette densité.

Pour ce faire, utilisons d'abord le fait que  $D(t)$  peut rejoindre et dépasser la résolution d'un élément capteur, auquel cas le signal observé reste à l'état de mélange. Partons du temps de mélange d'une salle acoustique, tel que défini par Polack [22], et qui correspond à l'instant où la densité des trajets devient supérieur la résolution auditive de l'oreille humaine :

$$t_m = \sqrt{\frac{10V}{0.056\pi c_0^3}} \approx 230 \text{ ms.} \quad (2.25)$$

Ce temps, s'il est particulièrement employé dans la communauté du traitement acoustique des salles, correspond également aux performances des microphones du commerce. De ma-



nière analogue à Polack, mais plus générale, définissons un instant où la densité de trajets dépasse la résolution d'acquisition donnée par le théorème de Shannon-Nyquist (la période d'échantillonnage étant  $1/f_s$ ) :

$$t_c = \sqrt{\frac{f_s V}{\pi c_0^3}} \approx 250 \text{ ms} \quad , \quad (f_s = 44.1 \text{ kHz}). \quad (2.26)$$

Finalement, suivons une autre manière de voir le problème et proposons un temps qui ne dépend que du milieu et de la source qui l'éclaire, puisqu'a priori le temps de début de la coda est une propriété intrinsèque au milieu de propagation. Considérons le temps caractéristique du milieu  $\tau = \sqrt[3]{V}/c_0$  et définissons une limite de résolution à l'échelle de ce temps. Si chaque trajet occupe un support temporel de longueur  $1/B$ , où  $B$  est la largeur de bande du champ, alors  $\tau B$  trajets peuvent cohabiter sans intersection sur une durée  $\tau$ .  $D(t)$  atteint cette valeur lorsque :

$$t_{\tau B} = \sqrt{\frac{\tau B V}{4\pi c_0^3}} \approx 120 \text{ ms}. \quad (2.27)$$

Ce dernier temps correspond au temps à partir duquel il y a un début d'interférence entre les trajets, et constitue une borne inférieure pour déterminer quand commence la coda acoustique. D'autres temps ont été définis dans la littérature, comme le temps de Ehrenfest pour des système chaotiques, le temps de Eisenberg et le temps de Thouless défini sur l'énergie perdue par diffusion, mais de manière générale, nous retiendrons ici que pour une réponse impulsionnelle acoustique dans une salle de type bureau meublé, le premier tiers de la réponse ne devrait pas être inclus dans la coda. Notons que c'est le RSB qui détermine où la coda s'arrête. Dans l'expérience présentée, un événement se termine lorsque le niveau d'énergie instantanée du signal passe en dessous de 10 fois l'énergie du bruit de mesure. Ce seuil a été choisi car il conduit à des estimations robustes et répétables tout en maximisant la longueur de la coda, et donc la quantité d'information à exploiter. Un seuil plus bas que 5 fois le niveau du bruit est la limite à laquelle le protocole est robuste ; en dessous, le calcul des corrélations sur des temps finis et courts n'est plus fiable, car la perturbation de la mesure est trop forte.

La corrélation de Green est estimée en inter-corrélant des extraits de codas, vues à chacun des capteurs. Le protocole est résumé dans l'Algo. 2. Le temps de propagation est 1.40 ms. Chacune des corrélations d'extraits de coda est calculée sur un support de 5 ms ( $5 \geq 1.4$ ), contraignant les données à traiter à une taille raisonnable : 5 ms à 44.1 kHz implique 221 échantillons. Les extraits de coda peuvent se superposer (« overlap » non nul), afin de :

- se servir de l'ensemble des données disponibles ;
- assurer la convergence des corrélations vers les composantes les plus corrélées.

Par ailleurs, afin de limiter les effets de bords induits par le calcul numérique de corrélation sur un temps fini, les extraits de codas se voient appliquer une fenêtre d'apodisation. La nature de la fenêtre, s'il s'agit d'une fenêtre classique de type Blackman par exemple, importe assez peu, dans la mesure où elle est assez « lisse » et dont l'auto-corrélation ne présente pas de lobes susceptibles de nuire à la détection du temps de propagation. Dans cette section, ce sont

250 corrélations qui sont cumulées pour former l'estimateur de la corrélation de Green. Une étude de convergence de cet estimateur sera menée par la suite.

---

**Algorithme 2:** Estimation de la corrélation de Green à partir de sources impulsionnelles.

---

- extrairement les évènements impulsionnels;	1
<b>forall the évènements do</b>	2
- trouver où la coda commence et s'arrête;	3
- extraire la coda $c_1$ du microphone 1;	4
- extraire la coda $c_2$ du microphone 2;	5
- extraire $N$ portions synchronisées de codas de longueur donnée $L$ dans $c_1$ et $c_2$ ;	6
<b>forall the paires d'extraits de codas do</b>	7
- calculer la corrélation entre $c_1$ et $c_2$ ;	8
- normaliser la corrélation obtenue;	9
- ajouter la corrélation obtenue aux autres déjà calculées;	10
<b>end</b>	11
<b>end</b>	12

---

Un résultat obtenu pour des signaux pleine bande est présenté dans la Fig. 2.6. La corrélation de la coda (en rouge), n'est pas modélisée (pour le moment) et ne peut donc pas servir pour estimer le temps de propagation  $t_1$ . C'est pour cette raison que l'identité de Ward est appliquée, afin de retrouver la fonction de Green (en bleu). C'est sa partie impaire qui est estimée.  $t_1$  est facilement localisé.

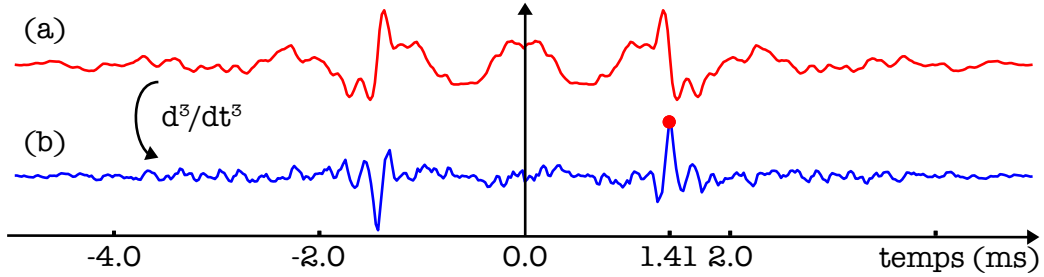


FIGURE 2.6 – Exemple type de résultat d'identification passive. (a) Corrélation de Green. (b) Partie impaire de la fonction de Green. Le temps de propagation extrait par argmax est localisé par le point rouge.

### 2.3.2 Analyses des performances expérimentales

Nous proposons maintenant d'évaluer les performances du protocole d'estimation introduit, face à des champs de bandes variables. Un banc de filtres passe-bas est appliqué sur les champs mesurés, avec pour fréquences de coupure  $\{2, 5, 9, 18, 22\}$  kHz. Les claquements de main excitent le milieu avec une énergie initiale uniformément répartie dans cette bande, voir le spectrogramme Fig. 2.7. Au cours de la propagation, la dissipation réduit de manière non négligeable la bande apparente du champ, en fonction du carré de la fréquence rappelons-le; la bande apparente étant définie comme la bande de fréquence où l'énergie du signal est

supérieure à celle du bruit d'acquisition.

### Légende (Figure 2.7)

Nous observons que :

- Dans les régions (a), le bruit de mesure domine et le RSB  $r_{dB}$  :

$$r_{dB} = 10 \log \left( \frac{\mathbb{E} [champ^2]}{\mathbb{E} [bruit d'acquisition^2]} \right)$$

est par conséquent très faible. Il est même tellement faible que la région est mise de côté et n'est pas utilisée par le protocole d'identification présenté ici.

- Dans la région (b), les trajets primaires qui atteignent le capteurs ne sont pas encore denses ; les signaux sont donc très cohérents. Les différents trajets, bien que sporadiques, sont corrélés entre eux mais ne répondent pas aux critères recherchés pour être utiles au protocole utilisant l'identité de Ward ;
- La région (coda) est identifiée comme exploitable pour appliquer l'identité de Ward.

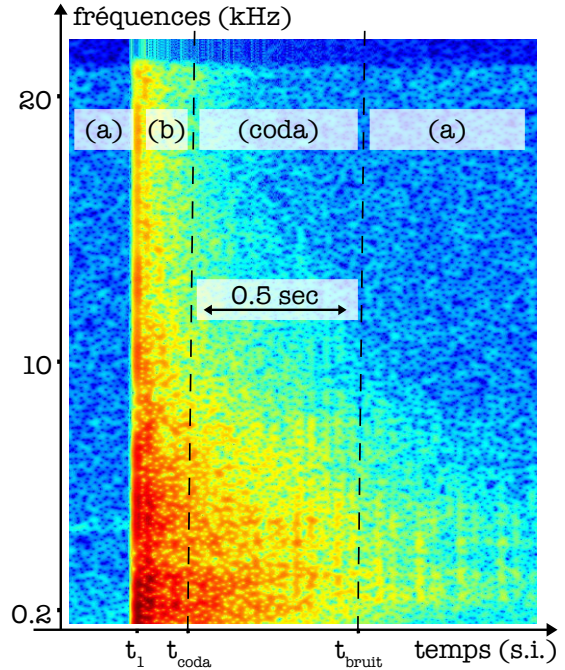


FIGURE 2.7 – Spectrogramme d'une impulsion.

Le moyennage des corrélations de codas prend alors tout son sens, afin de faire ressortir les éléments cohérents rémanents présents dans les champs mesurés, et qui correspondent ici à l'information sur le milieu de propagation.

Dans la Fig. 2.9, une performance de l'estimateur du temps de propagation est présentée en fonction de la bande des champs mesurés.

Dans la région (a) de la Fig. 2.9, l'augmentation de la bande-passante conduit à une réduction évidente du nombre d'erreurs de détection. En conséquence, l'écart-type de l'erreur d'estimation s'en retrouve considérablement amélioré. Ce résultat est en adéquation avec l'expression de la probabilité de l'erreur théorique de détection du temps propagation Eq. (2.20), fonction du rapport signal-à-bruit : le niveau de bruit est constant mais la quantité d'information utile augmente avec la bande, conduisant à une amélioration notable du RSB.

Dans la région (b) de la Fig. 2.9, la bande-passante varie de 7 à 15 kHz. La probabilité d'erreur de détection devient inférieure à 0.001, représentant de ce fait un évènement suffisamment improbable pour être négligé. Il s'avère que la borne Eq. (2.19) n'est pas atteinte (elle vaut par exemple  $10^{-6}$  ms à 9 kHz), car les limitations techniques bornent les performances. En effet, nous considérons deux faits. Premièrement, extraire l'argmax de la série temporelle conduit à un multiple de la période d'échantillonnage, à moins de déployer une technique plus évoluée d'estimation ; par exemple, une technique de régression polynomiale pour rechercher un maximum hors de la grille d'échantillonnage. Deuxièmement, les statistiques sont calculées sur un nombre fini de répétitions de l'expérience. La Fig. 2.10 montre que ce sont principa-

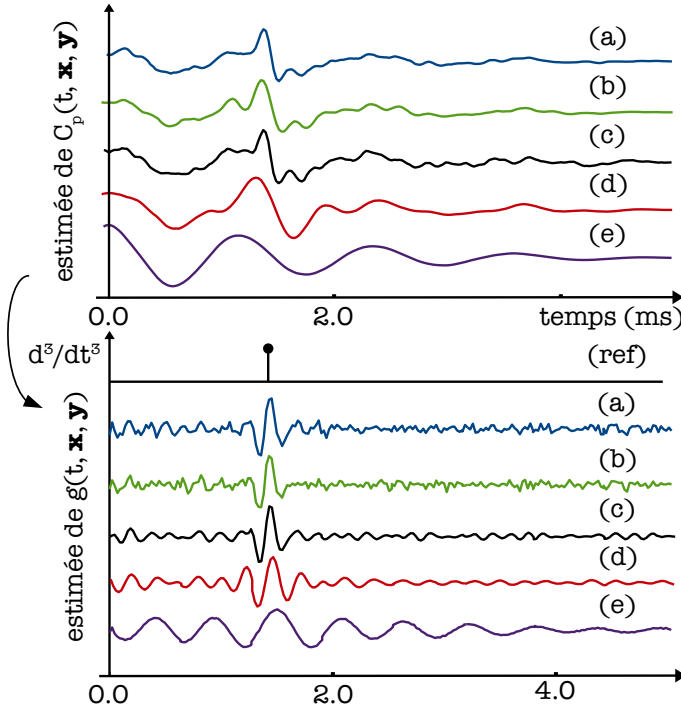


FIGURE 2.8 – Signaux estimés.

**Légende** (Figure 2.8)

Reconstructions typiques de la corrélation de Green et de la fonction de Green correspondante, aux différentes bandes étudiées.

- Haut : corrélations de Green estimées à partir de la corrélation d'extraits de codas. Les champs ont été filtrés passe-bas, avec pour fréquences de coupure
- 22 kHz (a) ;
- 18 kHz (b) ;
- 9 kHz (c) ;
- 5 kHz (d) ;
- 2 kHz (e) ;
- Bas : fonctions de Green estimées par application de l'identité de Ward sur les corrélations de Green. La fonction (ref) localise le temps de propagation à 1.40 ms.

**Légende** (Figure 2.9)

La figure présente l'écart-type du temps de propagation estimé  $\hat{t}_1$  en fonction de la bande passante des signaux enregistrés, estimé à partir de 1000 itérations du protocole.

- Pointillés : résolution due à l'échantillonnage et au nombre fini d'itérations utilisées pour calculer les statistiques,

$$9.7 \times 10^{-3} \text{ ms.}$$

- En gras : écart-type en utilisant tous les temps de propagation estimés ;
- Fin : écart-type calculés à partir des données excluant les erreurs de détection.

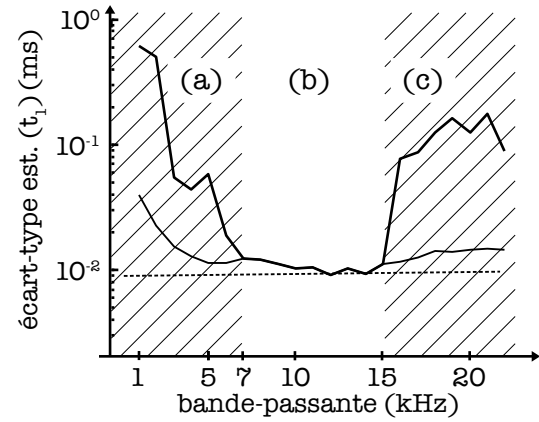


FIGURE 2.9 – Performances d'estimation.

lement deux localisations qui sont estimées pour le temps de propagation. Une formulation équivalente est de dire qu'estimer les temps autres que la position  $n$  et  $n + 1$  est moins probable qu'une probabilité seuil  $\epsilon$ , ou encore que : la probabilité que le temps de propagation soit compris dans un segment de longueur  $3 T_s$  est alors plus grande que  $1 - 2\epsilon$ .

Dans l'expérience,  $T_s = 1/44100$  s et  $\epsilon = 10^{-3}$  puisque ce sont 1000 itérations du protocole qui ont été réalisées pour estimer les statistiques. En appliquant la relation Eq. (2.28), l'écart-type de résolution  $\sigma$  est  $9.7 \times 10^{-3}$  ms, ce qui est en adéquation avec les résultats expérimentaux.

Si la densité de probabilité  $p_{t_1}$  du temps de propagation est modélisée par une distribution normale, de moyenne  $t_1$  et de variance  $\sigma^2$ , alors la relation suivante est vérifiée :

$$\operatorname{erf}\left(\frac{3T_s}{2\sqrt{2}\sigma}\right) \geq 1 - 2\epsilon. \quad (2.28)$$

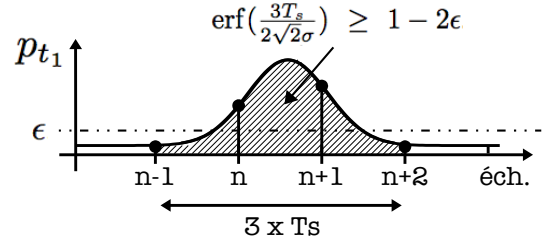


FIGURE 2.10 – Histogramme des estimations

Dans la région (c) de la Fig. 2.9, la bande passante est supérieure à 15 kHz. Hors de cette bande, la réponse du microphone s'écroule et le RSB augmente drastiquement. En conséquence, les erreurs de détection deviennent à nouveau non négligeable, conduisant à une baisse notable des performances. La différence notable entre les deux courbes est exclusivement due aux erreurs de détection, qui lorsqu'elles sont présentes, témoignent que le régime de fonctionnement (la bande passante) n'est pas adapté.

En résumé, il s'avère que les erreurs de détection sont responsables des plus grandes variations de performances. C'est pour une bande passante de 7 à 15 kHz que les erreurs de détection sont les moins probables et dans ces bandes, nous avons observé que c'est la résolution de l'acquisition qui limite les performances. Même ainsi, l'écart-type de la distance inter-capteur estimée est comparable aux dimensions des capteurs, soit de l'ordre du cm, ce qui est donc très bon. En réalité, même avec aussi peu de données utilisées (2 voies à 44.1 kHz, 0.5 secondes d'enregistrement, soit 88 ko de données), c'est la précision de l'acquisition qui est atteinte, montrant que l'algorithme converge très vite. Pour aller au-delà, des méthodes super-résolution d'extraction de l'argmax sont une piste à explorer.

## 2.4 Caractérisation des performances dans un milieu contrôlé

Jusqu'à présent, nous avons caractérisé les performances d'estimation du temps de propagation  $t_1$  et nous avons montré que ce dernier était estimé avec la précision de la chaîne d'acquisition. A présent, intéressons-nous aux fonctionnelles de Green. Sous couvert de trouver une base de représentation pour décrire les champs qui se propagent dans un milieu donné, il deviendrait possible de comparer les expressions analytiques des quantités d'intérêts (fonction/corrélation de Green) avec celles obtenues par l'expérimentation. Cette comparaison offrirait un retour sur le niveau de modélisation adopté et par extension, sur la pertinence des estimateurs proposés. C'est ce que nous nous apprêtons à aborder.

### 2.4.1 Identité de Ward en milieu fini, conditions de Neuman

Un mode  $\mathbf{e}(\mathbf{x})$  est un vecteur propre de l'opérateur responsable de la propagation du champ ; le Laplacien  $\Delta$  dans le cas de la pression acoustique. Il est équivalent de dire qu'un

mode est une solution spatiale à l'équation de propagation en régime libre, dont les caractéristiques ne dépendent que du milieu de propagation et pas de la source :

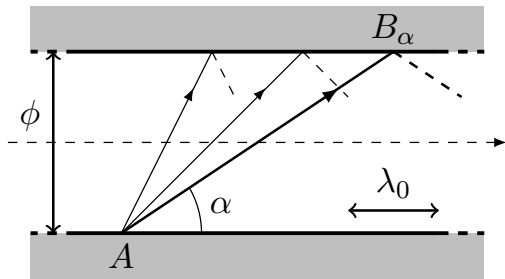
$$\Delta \mathbf{e}(\mathbf{x}) = -k^2 \mathbf{e}(\mathbf{x}), \text{ où } k^2 = \frac{\omega^2}{c_0^2 + j\alpha^2 \omega}. \quad (2.29)$$

L'ensemble des modes forme une base intrinsèque au milieu de propagation, utilisable pour représenter tout champ propagé. Par ailleurs, nous savons qu'un opérateur compact et borné (de bord suffisamment régulier, par exemple  $\mathcal{C}^\infty$  par morceaux) possède un spectre discret représenté par des ensembles de valeurs propres  $(k_n^2)$  et de vecteurs propres  $(\mathbf{e}(\mathbf{x})_n)$  associés dénombrables. C'est ainsi que nous pouvons donc caractériser les modes d'un milieu de propagation visco-acoustique confiné et les utiliser pour représenter tout champ capable de se propager dans ce milieu.

Pour illustrer la notion de mode et de vecteur d'onde, je propose de reprendre le concept du guide d'onde. Tout d'abord, notons que l'effet d'un obstacle sur la propagation d'une onde se traduit par un jeu de contraintes sur l'amplitude du champ, fonction du volume de l'obstacle, de sa position et de ses caractéristiques physiques (par exemple d'absorption). Pour un champ de pression acoustique, une interface isolante est modélisée par une condition de Neuman, c'est-à-dire l'annulation de la dérivée du champ au niveau de l'obstacle,

$$\left. \frac{\partial p(t, \mathbf{x})}{\partial t} \right|_{\mathbf{x} \in \partial M} = 0. \quad (2.30)$$

Par ailleurs, considérons une onde monochromatique (et l'extension du raisonnement à une onde de spectre quelconque s'appuiera sur le principe de superposition dans un milieu de propagation linéaire). L'onde est née d'un phénomène oscillatoire, de fréquence donnée  $f_0$ , comme le battement périodique d'une membrane de haut-parleur, et est transmise à l'air par couplage mécanique. Une fois dans l'air, l'onde se propage à une certaine vitesse  $c_0$ , ce qui conduit à une périodicité spatiale appelée longueur d'onde,  $\lambda_0 = c_0/f_0$ .



Considérons la propagation monodirectionnelle de cette onde monochromatique dans un guide infini. Les bords du guide déterminent les conditions de Neuman. Il apparaît que pour que la dérivée du champ soit nulle sur les bords, les distances  $AB_\alpha$  autorisées doivent être des multiples de  $\lambda_0$ ,

$$AB_{\alpha_n} = n\lambda_0. \quad (2.31)$$

L'angle d'incidence  $\alpha$  de l'onde est alors contraint à prendre un ensemble discret dénombrable de valeurs  $(\alpha_n)$ , telles que  $\cos(\alpha_n) = \phi/n\lambda_0$ . Ainsi, pour une source omnidirectionnelle monochromatique qui émettrait à une extrémité du guide, au multiple  $n$  de  $\lambda_0$  correspondrait un front d'onde en phase, car décomposable en un ensemble de rayons se propageant tous

avec des directions d'incidences  $\alpha_n$ . C'est sur ce principe que la fibre optique permet de transmettre de l'information sur de longues distances en restant structurée, moyennant une perte par propagation et par réflexion minimisée par le choix des matériaux et des dimensions.

Malheureusement, l'expression des modes d'un milieu de propagation est connue dans quelques cas seulement, où les milieux présentent des symétries fortes, comme pour un parallélépipède ou un cylindre. Par ailleurs, la transformation qui relie la géométrie du milieu aux modes n'est pas bijective ; un spectre peut caractériser plusieurs milieux et il n'est donc « pas possible d'entendre la forme d'un tambour », voir [24] pour en savoir davantage. Dans la suite, nous utiliserons néanmoins la représentation modale pour expliciter la corrélation de Green et la fonction de Green de manière générique. Une expression spécifique, et simple, d'un milieu de propagation étudié expérimentalement servira ensuite pour confronter les grandeurs physiques expérimentalement obtenues à leur pendant théorique.

#### 2.4.1.1 Fonction de Green et base de représentation

Rappelons que la fonction de Green est la solution de l'équation de propagation lorsque le milieu est excité par une impulsion. La transformée de Fourier de l'équation de propagation (B.1), lorsque la source est une impulsion, s'écrit :

$$[-\omega^2 + j\omega\mathcal{D} + \mathcal{L}]g(\omega, \mathbf{x}, \mathbf{y}) = \delta_{\mathbf{x}, \mathbf{y}} \quad (2.32)$$

où  $\mathcal{D}$  est la transformée de Fourier de l'opérateur de dissipation et  $\mathcal{L}$  est la transformée de Fourier du Laplacien. Les opérateurs  $\mathcal{D}$  et  $\mathcal{L}$  sont auto-adjoints, voir [25] pour davantage d'explications ; il existe une base  $(\mathbf{e}_n)_{n=1\dots N}$  qui permet de les décomposer simultanément. Ceci conduit aux relations suivantes :

$$\begin{cases} \langle [\mathcal{L}]g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle &= \langle g(\omega, \mathbf{x}, \mathbf{y}), [\mathcal{L}]\mathbf{e}_n \rangle \\ \langle [\mathcal{D}]g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle &= \langle g(\omega, \mathbf{x}, \mathbf{y}), [\mathcal{D}]\mathbf{e}_n \rangle \end{cases} \quad (2.33)$$

et par conséquent, la projection de l'Eq. (2.32) sur la base  $(\mathbf{e}_n)_{n=1\dots N}$  s'exprime

$$\begin{aligned} -\omega^2 \langle g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n(\mathbf{x}) \rangle &= -\omega^2 \langle g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n(\mathbf{x}) \rangle \\ + i\omega \langle [\mathcal{D}]g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n(\mathbf{x}) \rangle &\Leftrightarrow + i\omega \langle g(\omega, \mathbf{x}, \mathbf{y}), [\mathcal{D}]\mathbf{e}_n(\mathbf{x}) \rangle \\ + \langle [\mathcal{L}]g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n(\mathbf{x}) \rangle &+ \langle g(\omega, \mathbf{x}, \mathbf{y}), [\mathcal{L}]\mathbf{e}_n(\mathbf{x}) \rangle \\ = \langle \delta_{\mathbf{x}, \mathbf{y}}, \mathbf{e}_n(\mathbf{x}) \rangle &= \mathbf{e}_n(\mathbf{y}) \end{aligned} \quad (2.34)$$

Dans la base  $(\mathbf{e}_n)_{n=1\dots M}$ , notons  $(l_n)_{n=1\dots N}$  et  $(d_n)_{n=1\dots N}$  les valeurs propres des opérateurs  $\mathcal{L}$  et  $\mathcal{D}$  respectivement, telles que par définition :

$$\begin{cases} \mathcal{L}\mathbf{e}_n &= l_n\mathbf{e}_n \\ \mathcal{D}\mathbf{e}_n &= d_n\mathbf{e}_n \end{cases}, \quad \forall n = 1 \dots N. \quad (2.35)$$

Ce faisant, la projection de la fonction de Green sur chaque vecteur  $\mathbf{e}_n$  de la base s'écrit :

$$\langle g(\omega, \mathbf{x}, \mathbf{y}), \mathbf{e}_n(\mathbf{x}) \rangle = \frac{\mathbf{e}_n(\mathbf{y})}{-\omega^2 + j\omega d_n + l_n} = -\frac{\mathbf{e}_n(\mathbf{y})}{(\omega - \omega_n^+)(\omega - \omega_n^-)} \quad (2.36)$$

où  $\omega_n^\pm = jd_n \pm \omega_{\alpha,n}$ , avec  $\omega_{\alpha,n} = \sqrt{c_0^2 l_n^2 - d_n^2}$ .  $\alpha$  est coefficient d'amortissement dans l'opérateur de dissipation  $\mathcal{D}$  et  $c_0$  la vitesse de propagation.

L'expression de la fonction de Green s'obtient en considérant toutes les projections :

$$g(\omega, \mathbf{x}, \mathbf{y}) = - \sum_{n \geq 0} \frac{\mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y})}{(\omega - \omega_n^+)(\omega - \omega_n^-)} \quad (2.37)$$

Finalement, pour basculer dans le domaine  $(t, \mathbf{x})$ , nous proposons d'utiliser le théorème des résidus (preuve fournie dans l'Annexe B.5) :

$$g(t, \mathbf{x}, \mathbf{y}) = u(t) \sum_{n \geq 0} \frac{\exp\left(-\frac{d_n}{2}t\right) \sin(\omega_{\alpha,n}t)}{\omega_{\alpha,n}} \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y}). \quad (2.38)$$

Si la dissipation est constante, il est alors facile de se ramener à l'expression monochromatique de la réponse impulsionnelle de l'oscillateur amorti, où le terme d'amortissement est contenu dans l'exponentielle et intervient également dans la pulsation du circuit :

$$h(t) = \frac{1}{L\omega} \exp(-\eta t) \sin(\omega t) \times u(t).$$

#### 2.4.1.2 Corrélation de Green en milieu borné

A partir de l'Eq. (2.37), l'expression dans le domaine fréquence-espace de la corrélation de Green dans la base étudiée est :

$$C_g(\omega, \mathbf{x}, \mathbf{y}) = \sum_{n \geq 0} \frac{\mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y})}{|\omega - \omega_n^+|^2 |\omega - \omega_n^-|^2} \quad (2.39)$$

d'où à nouveau, une expression dans le domaine temps-fréquence peut être déduite par transformée de Fourier inverse (la preuve est présentée dans l'Annexe B.6) :

$$C_g(t, \mathbf{x}, \mathbf{y}) = \sum_{n \geq 0} \frac{\exp\left(-\frac{d_n}{2}|t|\right) (\omega_{\alpha,n} \cos(\omega_{\alpha,n}t) + d_n \sin(\omega_{\alpha,n}|t|))}{\omega_{\alpha,n} l_n^2 d_n} \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y}), \quad (2.40)$$

A ce stade, notons que les pôles de la corrélation de Green sont les pulsations modales du milieu de propagation  $\omega_n^\pm$ . C'est un résultat intéressant puisque cela signifie qu'en pratique, l'estimation de la corrélation de Green suffit pour estimer les modes du milieu de propagation. Le spectre de l'opérateur de propagation ne nécessite donc pas l'application de l'identité de Ward pour être identifié. En pratique, ce principe est appliqué pour la surveillance de l'état de santé d'une structure, voir par exemple l'article [26] qui fait référence dans le domaine et qui exploite l'équation discrète de Navier pour les solides élastiques afin de modéliser un pont. Une technique de SHM qui en découle, le « mode tracking », consiste à suivre l'évolution de ce spectre dans le temps, puisqu'il est une image de la géométrie de la structure. C'est un exemple de paramètre qui ne souffre pas de la propagation d'erreur induite par l'application de l'identité de Ward pour être estimable. Notons qu'à partir des expressions de la fonction de Green (2.38) et de la corrélation de Green (2.40), l'identité de Ward est retrouvée en appliquant l'approximation de faible viscosité  $\omega_{\alpha,n} \approx c_0^2 l_n$ .



### 2.4.2 Application : cas du parallélépipède rectangle

Exploitions à présent l'expression analytique obtenue des quantités de Green. Nous nous intéressons dans cette section au cas du parallélépipède rectangle de dimensions  $L_x \times L_y \times L_z$ . Les paramètres du modèle modal sont connus et valent :

$$\begin{aligned} l_n^2 &= \pi^2((n_x/L_x)^2 + (n_y/L_y)^2 + (n_z/L_z)^2), \\ d_n &= \alpha^2 l_n^2 / 2, \\ \omega_{\alpha,n} &= \sqrt{c_0^2 l_n^2 - d_n^2}, \\ \mathbf{e}_n(\mathbf{x}) &= L_x L_y L_z \cos(\pi n_x \mathbf{x}_1 / L_x) \cos(\pi n_y \mathbf{x}_2 / L_y) \cos(\pi n_z \mathbf{x}_3 / L_z) / 8. \end{aligned} \quad (2.41)$$

#### Légende (Figure 2.12)

Le milieu étudié est un parallélépipède de dimensions  $L_x = 0.8\text{m}$ ,  $L_y = 1.2\text{m}$  et  $L_z = 0.6\text{m}$ . Dans la Fig. 2.12, un des panneaux latéraux de la boîte est laissé ouvert, pour permettre d'observer la configuration de l'instrumentation à l'intérieur. Le panneau était fermé pendant les manipulations. Deux microphones omnidirectionnels prosignal sont positionnés en :

$$\begin{aligned} \mathbf{A} &= \{L_x/2, L_y/4, L_z/2\} \\ \mathbf{B} &= \{L_x/2, 3L_y/4, L_z/2\}. \end{aligned}$$

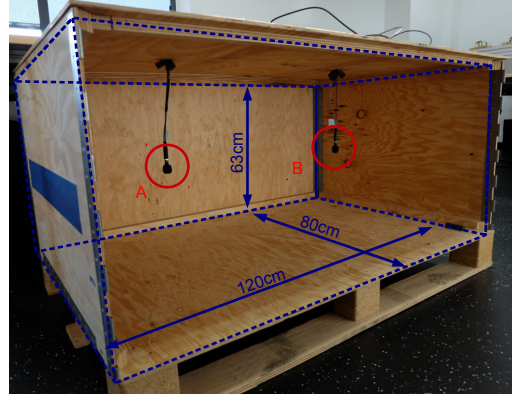


FIGURE 2.12 – Configuration expérimentale

#### 2.4.2.1 Vérification de l'identité de Ward par la simulation

L'approximation réalisée par l'identité de Ward est quantifiée par simulation, à partir des expressions théoriques des quantités de Green. La Fig. 2.13 permet tout d'abord de visualiser la fonction de Green et la corrélation de Green correspondant à la configuration présentée Fig. 2.12. La fonction de Green est confrontée à celle obtenue par application de l'identité de Ward à la corrélation de Green théorique ; une mesure de leur ressemblance est proposée avec l'écart quadratique moyen normalisé entre les deux fonctionnelles, évalué à

$$\frac{\sum_{k=1}^K \left( g_{th}[k] - \frac{\partial^3}{\partial t^3} c g_{th}[k] \right)^2}{\sum_{k=1}^K g_{th}[k]^2} = 1\%. \quad (2.42)$$

L'approximation commise par l'hypothèse de faible viscosité est faible ; l'approximation de Ward pour un milieu faiblement visqueux est donc exploitable, en particulier pour l'estimation de paramètres comme le temps de propagation  $t_1$ .

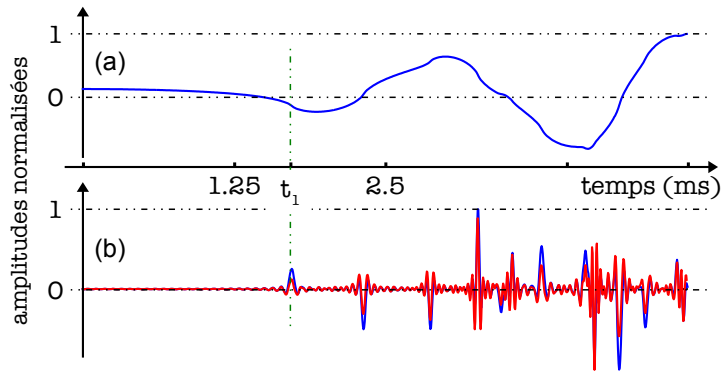


FIGURE 2.13 – Fonctionnelles de Green.

**Légende** (Figure 2.14)

Quantités de Green construites à partir de leurs expressions modales Eq. (2.40) et Eq. (2.38), sur un horizon temporel de 5 ms et un support spectral  $[0-40]$  kHz :

- (a) Corrélation de Green ;
- (b) Fonction de Green théorique et estimée par application de l'identité de Ward sur la corrélation (a).

#### 2.4.2.2 Influence de la largeur de spectre

A présent, observons les différentes formes que prennent les quantités de Green lorsque la bande passante des signaux (c'est-à-dire le nombre de modes utilisés dans la modélisation) varie, afin d'en étudier l'influence sur l'estimation de paramètres, voir Fig. 2.14.

**Légende** (Figure 2.14)

Les quantités de Green sont construites à partir de leur expression modale, avec une bande passante différente :

- (a) Corrélations de Green
- (b) Fonctions de Green.

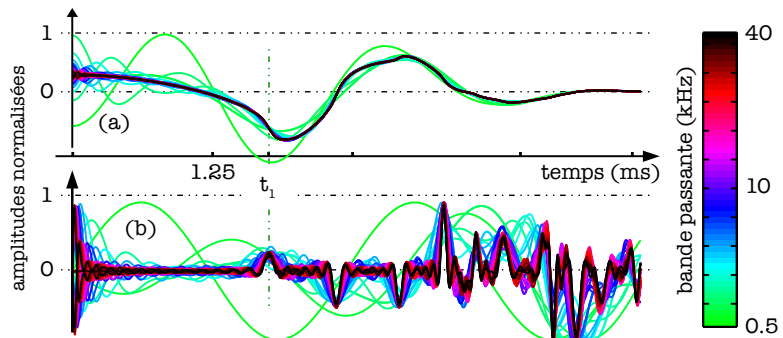


FIGURE 2.14 – Signaux à bande limitée

La corrélation de Green varie moins que la fonction de Green lorsque sa bande passante change. Dans les plus basses fréquences, un fort caractère oscillant se démarque vers zéro (temps nul), réhibitoire dans la fonction de Green pour l'estimation de paramètres comme le temps de propagation. Deuxièmement, la largeur des temps de propagation dépend de la bande (du nombre de modes considérés). C'est un problème d'estimation classique où la courbure du lobe peut être reliée à la variance de l'estimateur du temps de propagation.

#### 2.4.2.3 Comparaison des résultats théoriques et expérimentaux

Nous cherchons à présent à comparer l'expression théorique des quantités de Green à celles obtenues par l'expérience. Le protocole d'identification expérimental de la section précédente est appliqué dans un milieu différent. La configuration de l'expérience est représentée Fig. 2.15.

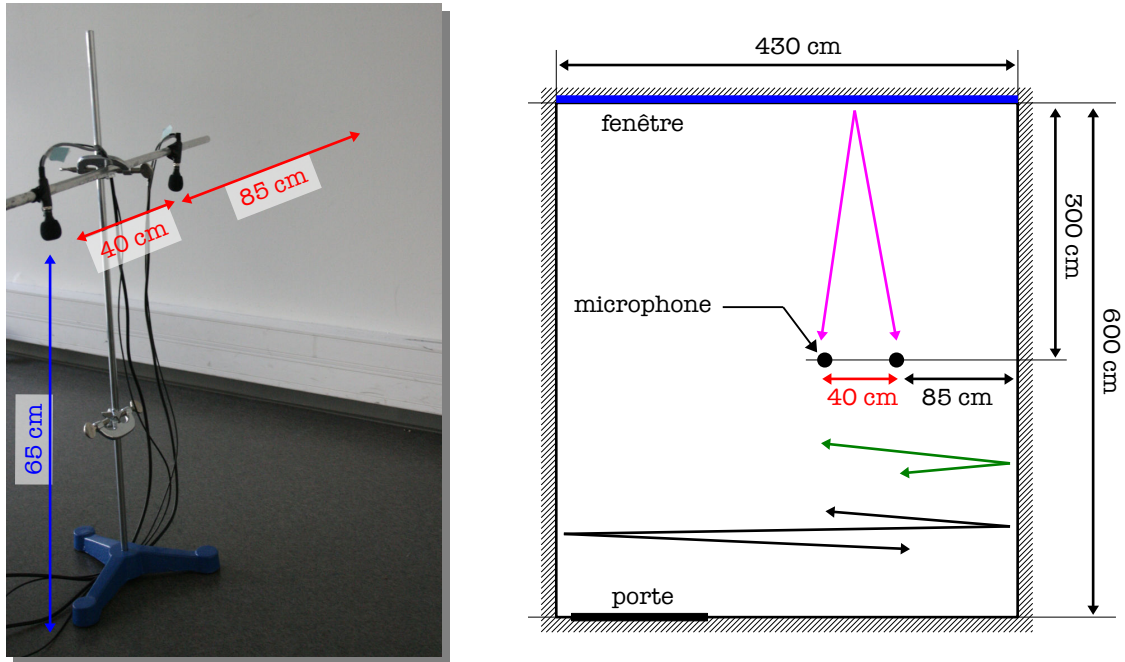


FIGURE 2.15 – Salle non meublée. Le trajet le plus court entre les deux microphones est en rouge et mesure 40 cm. Les trajets secondaires sont représentés, par ordre croissant de longueur, en vert, rose et noir.

Les évènements d'opportunité sont des claquements de mains répartis aléatoirement dans le volume de la pièce. Les quantités d'intérêts sont estimées en cumulant l'information apportée par une vingtaine d'évènements impulsionnels acoustiques, d'où quelques dizaines de milliers de segments de coda sont extraits. Ces segments mesurent 0.05 sec et conduisent à reconstruire les fonctions d'intérêt sur des distances parcourues par le front d'onde d'environ 17 m.

Dans la Fig. 2.16, la fonction de Green théorique (expression modale), utilisée comme référence, est comparée à sa version estimée.

A partir des concordances tirées entre les trajets, premièrement, il est clair que le trajet direct de la réponse impulsionnelle, qui correspond à la distance entre les deux capteurs, est correctement identifié. C'est la composante de la réponse qui présente le plus d'énergie puisqu'elle est la moins atténuée par la propagation. Ensuite, le trajet qui relie les deux microphones en rebondissant sur le sol n'est presque pas estimé, tandis que le suivant, correspondant au trajet vert sur la Fig. 2.15 l'est parfaitement. Ce résultat est attribué au fait que les murs (en béton) présentent un coefficient de réflexion bien supérieur à celui du sol (en lino). Ce phénomène est de nouveau observé avec le trajet supposé être réfléchi au plafond (faux-plafond en aggloméré) et d'autres trajets plus complexes à déterminer. Finalement, d'autres trajets sont identifiés : les trajets magenta et noirs sur la Fig. 2.15. Ces derniers, bien que composé de multiples réflexions, les ont réalisées sur les murs en béton exclusivement, préservant au maximum leur énergie. C'est le cas également pour 2 autres trajets, liés cette fois-ci à des trajectoires plus complexes.

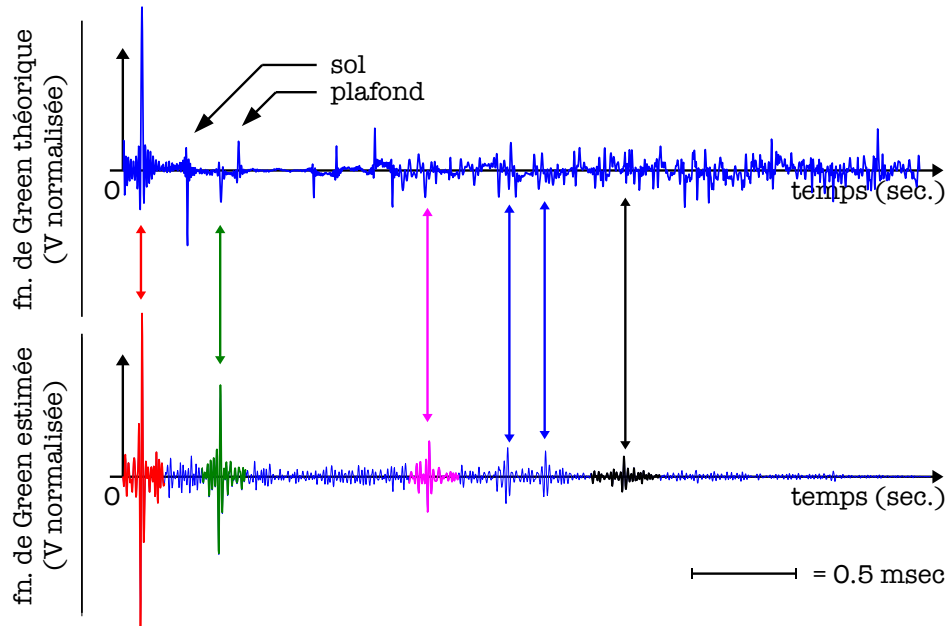


FIGURE 2.16 – Fonctions de Green théorique et estimée.

La reconstruction de la fonction de Green a été présentée sur un horizon temporel grand (environ 3 fois plus grand que la longueur caractéristique du milieu). Ce faisant, plusieurs trajets entre les deux capteurs ont été identifiés, correspondant parfaitement à la géométrie étudiée (erreur centimétrique). Cette approche a mis en évidence que l'absorption des murs n'était pas modélisée et qu'elle peut par ailleurs être estimée expérimentalement par cette approche. Il s'agirait par exemple de rendre de l'atténuation d'amplitude observée aux temps de propagation estimés contre celles attendues en théorie.

De manière un peu plus générale, la précision des estimateurs mis en oeuvre est déjà très bonne puisque de l'ordre de la taille des capteurs, mais le rapport signal à bruit du signal peut être amélioré en employant des pré/post-traitements inspirés des techniques en acoustique des salles (égalisation, conditionnement). Leur diversité et leur emploi étant conditionné aux exigences de performances attendues, une investigation sérieuse et poussée ne peut être menée qu'en répondant à un cahier des charges donné. En revanche, nous concluons cette section en exposant la corrélation de Green estimée, afin de montrer qu'elle ne correspond pas au modèle théorique, voir Fig. 2.17 ; ce résultat montre que les techniques d'égalisation utilisées ne sont peut-être pas suffisamment performantes ou que la modélisation de la physique n'est peut-être pas suffisamment complète. A l'heure actuelle, la corrélation de Green estimée reste une version filtrée de la corrélation vraie. La comparaison théorie/expérimentation ne peut donc pas être menée entièrement. En revanche, malgré ce filtrage, l'application de l'identité de Ward produit une fonction de Green exploitable pour estimer plusieurs temps de propagation ! Des post-traitements empruntés aux techniques d'acoustique des salles et de traitements de la musique peuvent être utilisés pour améliorer les résultats obtenus. Par exemple, ici, la phase dans la fonction de Green estimée est difficile à identifier. Or, cette phase indique, par application des lois de Snell-Descartes, la parité du nombre de réflexions vécues durant le trajet du front

d'onde. C'est donc une information qui peut être importante pour caractériser la géométrie du milieu. Pour la retrouver et même améliorer la réponse dans son ensemble, le filtrage de Kirkeby est le filtrage, s'il ne doit y en avoir qu'un seul, qui me semble être le plus adapté. Ce dernier est le filtre adapté à la réponse impulsionnelle de la chaîne d'acquisition, couplée à une règle de régularisation sur le spectre. Soit  $H(f)$  la réponse fréquentielle du filtre à blanchir, alors le filtre de Kirkeby s'écrit :

$$K(f) = \frac{H(f)^*}{|H(f)|^2 + \epsilon(f)} \quad (2.43)$$

où  $\epsilon(f)$  est la contrainte de régularisation du spectre. Dans la méthode suggérée, cette « pincée de bruit blanc » peut être choisie de manière à avoir un caractère sélectif sur la bande passante du signal à filtrer.

Généralement employé à l'acquisition pour enregistrer une réponse impulsionnelle la plus blanche possible, nous pouvons également l'appliquer à la fonction de Green estimée, corrigeant par là-même les distortions induites par l'application de l'identité de Ward et la corrélation des champs. La conséquence de ce post-traitement est l'amélioration du rapport signal à bruit, à l'instar d'un filtrage adapté. Davantage de traitements pourraient être appliqués pour améliorer la qualité des quantités estimées, mais étant donné la chaîne de mesure utilisée, le premier pas consisterait à utiliser des capteurs de meilleure qualité (directivité, réponse).

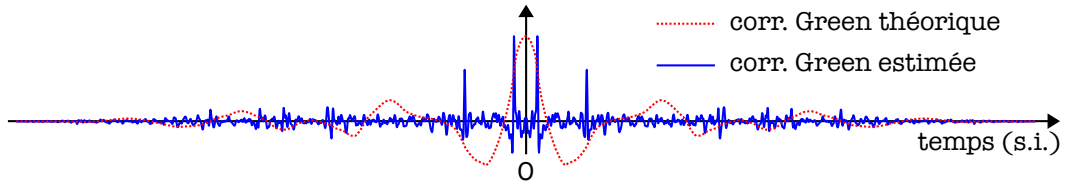


FIGURE 2.17 – Corrélation de Green théorique et estimée.

#### 2.4.2.4 Application dans un milieu avec changement d'échelle

La démarche de la section précédente est appliquée au milieu parallélépipédique (la boîte) dont la longueur caractéristique est plus de 5 fois plus petite que la salle de la manipulation précédente, voir Fig. 2.12. Des frappes sur la caisse font office de sources ponctuelles impulsionnelles aléatoires. Ici, c'est une étude de convergence des estimateurs de Green qui est présentée Fig. 2.18, convergence qualifiée en fonction de la quantité de données utilisées.

A partir des données expérimentales issues de l'expérience dans le parallélépipède, les fonctionnelles de Green converge, et elles le font avec peu de données. L'approche est robuste. En revanche, les formes d'onde théoriques et expérimentales ne coïncident pas. Pour la corrélation de Green, c'est un résultat déjà observé dans la section 2.4.2.3, où plusieurs temps de propagation sont malgré tout correctement estimés.

Ici, le changement d'échelle du milieu (d'une pièce à une boîte) conduit la source utilisée à mettre le milieu en résonance, permettant à certains modes de se démarquer des autres par une

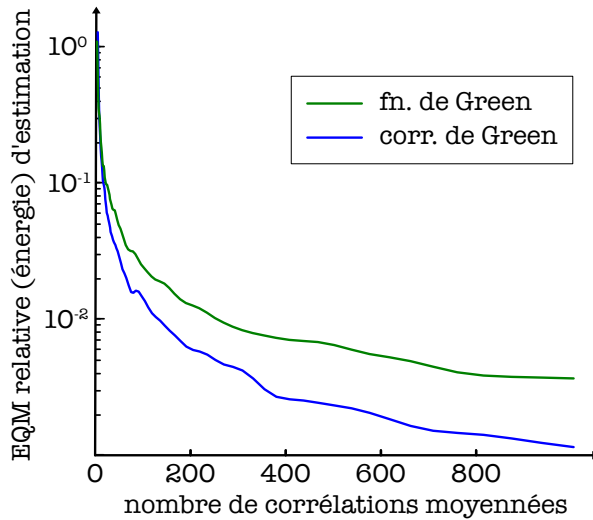


FIGURE 2.18 – Stabilité de la convergence

**Légende** (Figure 2.18)

Représentation d'un état de convergence de l'estimation des quantités de Green; une mesure de ressemblance (EQM relative à l'énergie de la cible) est présentée en fonction du nombre de corrélations de codas utilisées pour réaliser l'estimation. Les statistiques sont calculées à partir de 1000 réalisations du protocole d'identification.

**Légende** (Figure 2.19)

Spectrogramme d'une réponse impulsionnelle dans le milieu parallélépipédique :

- Dans les régions (n), le bruit de mesure prévaut sur le champ acoustique, dans le sens où son utilisation nuit à l'estimation de la corrélation de Green : ces régions ne sont pas utilisées ;
- La région (coda) débute lorsque  $t_1$  est perçu et se termine au temps de mélange. Le signal est composé de trajets résolus et donc fortement corrélés. Cette région n'est pas utilisée dans l'algorithme d'identification passive ;
- La coda se trouve au dessus de 2.5 kHz ;
- En-dessous de cette fréquence, région (m), la source a mis le milieu en résonance et le signal n'est pas blanc.

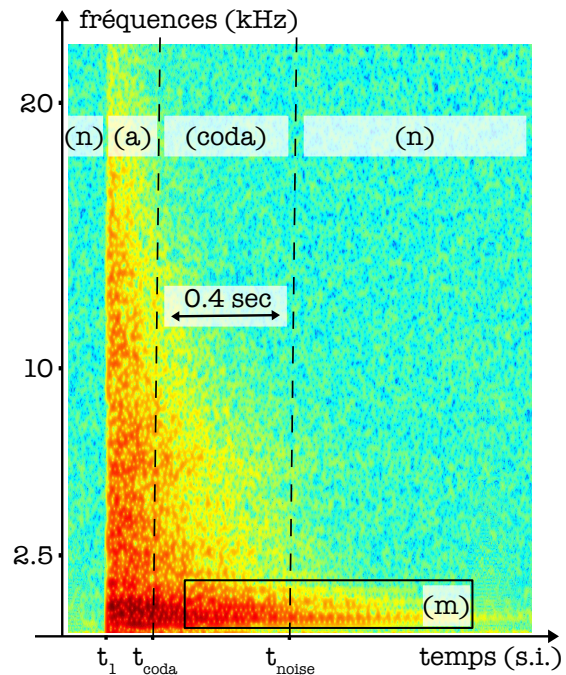


FIGURE 2.19 – Spectrogram d'une R.I.

énergie supérieure et récurrente, voir Fig. 2.19. Stricto sensu, il n'y a pas d'équipartition de l'énergie à un RSB suffisamment grand pour être exploitable par le protocole d'estimation. La corrélation de la partie réverbérée est une estimation de la corrélation de Green polluée par ces modes résonants à forte énergie, qui occultent l'information apportée par le milieu sur le reste du spectre. Cette configuration ne permet donc pas d'estimer la fonction de Green du milieu ou tout autre paramètre qui en découle. En conclusion, dans ces conditions expérimentales, la partie réverbérée du champ acoustique conduit à une corrélation de champ qui est estimée



de manière robuste. En revanche, cette corrélation n'est pas la corrélation de Green ; c'en est une version filtrée. Dans le cas présent, l'information trop inégale dans certaines bandes de fréquences ne permet pas de reconstruire la fonction de Green de manière exploitable (par exemple pour estimer le temps de propagation inter-capteur).

## 2.5 Discussion à propos de la distribution spatiale des sources

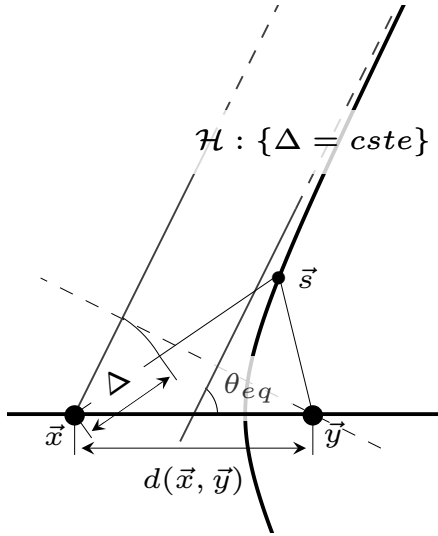
Dans cette dernière section, nous ouvrons la discussion sur les performances d'estimation des fonctionnelles de Green jusqu'alors tenue en regard de leur largeur spectrale en temps. L'objectif est de proposer un modèle qui puisse rendre compte de certaines des propriétés spatiales du champ mesuré. Pour simplifier la formalisation du problème, l'amplitude des champs sera normalisée et leur contenu spectral fixé une fois pour toute. Nous pourrions ainsi nous concentrer exclusivement sur les propriétés spatiales du champ. Le genre d'analyse proposée n'est pas sans antécédent dans la littérature. Entre autre, nous pouvons citer les travaux de Sanchez-Sesma qui se sert d'une représentation similaire pour interpréter des résultats expérimentaux. Ici, nous proposons un paradigme de représentation théorique qui nous mènera jusqu'à une borne de Cramér-Rao intégrant à la fois les propriétés temporelles et spatiales des champs mesurés.

### 2.5.1 Modèle géométrique de sources d'opportunité

Nous proposons un modèle de propagation 2D en milieu infini. Une onde se propage alors sans rencontrer d'obstacle et sur une distance qui dépend de sa puissance à l'émission et de l'atténuation géométrique et dissipative. Dans le contexte d'estimation passive, la position de la source, entre autre, est *a priori* non connue et non contrôlée. Cependant, les positions d'un ensemble de sources sont une propriété du champ susceptible de fortement influencer les performances d'estimation des fonctionnelles de Green (et par conséquent du temps de propagation inter-capteur). Nous sommes donc à la recherche du paramètre idéal pour paramétrer exhaustivement la distribution des sources, autre que leurs positions puisque celles-ci sont inconnues.

Considérons une source ponctuelle en  $\mathbf{s}$ , émettant un champ d'amplitude aléatoire, se propageant sur un front d'onde sphérique. Les deux capteurs en  $\mathbf{x}$  et  $\mathbf{y}$  perçoivent le front d'onde avec un décalage temporel, un retard du à la géométrie source-capteurs appelé différence de marche. Si le retard de propagation d'un front d'onde perçu par une paire de capteurs correspond à une différence de marche (c'est-à-dire que le retard de propagation est inférieur au temps de propagation inter-capteur), alors la source ayant généré le front d'onde était localisée sur des hyperboles dont les foyers sont les capteurs. La preuve de cette assertion géométrique est donnée dans l'Annexe B.7.

La combinaison de cette représentation avec le principe de Huygens-Fresnel offre un outil de modélisation très intéressant : nous pouvons inférer n'importe quel champ vu aux capteurs grâce à un ensemble de sources d'ondes planes. Ce faisant, les propriétés spatiales des sources

**Légende (Figure 2.20)**

Toute source ponctuelle qui induit une différence de marche  $\Delta$  vue aux capteurs est localisée sur une même hyperbole  $\mathcal{H}$ . A chaque  $\Delta$  correspond une unique source d'onde plane induisant cette différence de marche : son angle d'incidence  $\theta_{eq}$  sur les capteurs est obtenu en projetant la source à l'infini sur l'asymptote de l'hyperbole.

FIGURE 2.20 – Equivalence de sources

sont simples et interprétables : leur distribution et leurs statistiques peuvent être paramétrées par leur angle d'incidence sur les capteurs uniquement.

Considérons un ensemble de sources d'opportunité modélisées par un ensemble de sources d'onde plane. Ces sources sont indépendantes et identiquement distribuées, en temps comme en espace. Cela signifie que l'inter-corrélation des champs émis par des sources différentes est nulle, tandis que l'auto-corrélation normalisées des champs émis par chaque source est identique et vaut  $C_f(t)$ . Le bruit d'excitation du milieu est donc blanc spatialement. L'auto-corrélation des sources est normalisée afin de se séparer de l'effet d'atténuation due à la propagation ainsi que la puissance avec laquelle le champ est reçu aux capteurs. Enfin, les angles d'incidence de chaque source d'onde plane sont identiquement distribués selon la loi  $p_\theta$ . Appelons  $p(t, \mathbf{x}|i)$  la contribution du champ mesuré en  $\mathbf{x}$  qui provient de la source  $i$ , d'angle d'incidence  $\theta_i$  et perçue aux capteurs avec la puissance  $P_i$ . L'inter-corrélation des champs mesurés en  $\mathbf{x}$  et  $\mathbf{y}$ , composés de  $N$  source i.i.d. comme définies ci-dessus, s'écrit :

$$\mathbb{E}\left[\sum_{i=1}^N p(t, \mathbf{x}|i) \sum_{j=1}^N p(t+u, \mathbf{y}|j)\right] = \sum_{k=1}^N C_f(t - \Delta_k(\theta_k)). \quad (2.44)$$

**2.5.2 Modèle d'observation pour la corrélation de Green**

En utilisant des sources d'opportunité, les champs tels que les codas sont vus comme des processus stochastiques dont les statistiques permettent d'accéder à la corrélation de Green du milieu de propagation. Considérons dans cette section que les champs mesurés ont été normalisés pour tous avoir la même énergie. C'est une manipulation des champs similaire à la normalisation de la corrélation des segments de coda, présentée dans l'algorithme d'identification passive. Le dessein de cette mise à l'échelle est de s'abstraire des phénomènes d'atté-



uation d'amplitude due à la propagation. Pour le modèle proposé, la corrélation Eq. (2.44) est fondamentale pour obtenir un modèle d'observation de la corrélation de Green :

$$z(t) = \int_{-1}^1 p_{\Delta}(u) C_f(t - u) du + n(t), \quad (2.45)$$

ou de manière équivalente, en posant  $u = \cos(\theta)$ ,

$$z(t) = \int_{-\pi}^{+\pi} p_{\theta}(\theta) C_f(t - t_1 \cos(\theta)) \sin(\theta) d\theta + n(t). \quad (2.46)$$

où  $p_{\theta}$  et  $p_{\Delta}$  sont les densités de probabilité des angles d'incidence et des différences de marche, respectivement.

Dans l'Eq. (2.45), le résidu d'estimation  $n(t)$  est issu du calcul de corrélations à partir de champs quantifiés, à temps fini et entachés d'un bruit de mesure modélisé blanc additif et gaussien. Ce résidu peut être modélisé centré, de variance  $\sigma^2$ , voir [9] pour des justifications.

Indiquons que le modèle proposé est valide dans le sens où l'identité de Ward peut être vérifiée. Pour ce faire, faisons l'unique hypothèse que la distribution des différences de marche est uniforme entre  $-1$  et  $+1$ , auquel cas le modèle d'observation s'écrit

$$z(t) = [C_f * \Pi](t) + n(t) \quad (2.47)$$

où  $\Pi$  est la fonction porte :  $\Pi(-1 \leq t \leq 1) = 1$ , 0 ailleurs. Sachant qu'une porte est construite en combinant deux échelons et qu'un échelon est l'intégrale d'un dirac (une impulsion), alors la dérivation première de l'Eq. (2.47) amène à l'identité suivante :

$$\frac{\partial}{\partial t} z(t) = [C_f(t) * (\delta(t + t_1) - \delta(t - t_1))](t) + \frac{\partial n(t)}{\partial t} \quad (2.48)$$

et ainsi pour les temps positifs :

$$\frac{\partial}{\partial t} z(t)(t \geq 0) = -[C_f(t) * \delta(t - t_1)](t) + \frac{\partial n(t)}{\partial t}. \quad (2.49)$$

Dans l'Eq. (2.49), l'identité de Ward est reconnaissable dans le cas de la dissipation constante, retrouvée par exemple par J.-L. Lacoume [3], à un facteur multiplicatif près, perdu à cause de la normalisation des champs.

Le point important à ce stade est de remarquer que l'Eq. (2.45) nous donne un modèle d'observation de la corrélation de Green, de la forme classique où la nuisance est un bruit blanc additif gaussien. L'information de Fisher pour le temps de propagation  $t_1$ , notée  $F_{t_1}$ , s'écrit alors :

$$F_{t_1}(p_{\theta}) = \frac{1}{2\sigma^2} \int_{\mathbb{R}} \left[ \int_0^{2\pi} p_{\theta}(\theta) \cos(\theta) \frac{\partial}{\partial t} C_f(t - t_1 \cos(\theta)) d\theta \right]^2 dt.$$

Ainsi, la borne de Cramer-Rao présentée dans un cadre d'estimation active est ici applicable dans le contexte passif. Comme exposée par exemple dans [9], cette borne se déduit simplement

de l'information de Fisher, telle que

$$\mathbb{E}[(\hat{t}_1 - t_1)^2] \geq \frac{1}{F_{t_1}(p_\theta)}. \quad (2.50)$$

Finalement, pour résumer, le paradigme de représentation proposé permet de vérifier l'identité de Ward Eq. (2.48), retrouvée grâce à l'hypothèse de loi uniforme pour la distribution des différences de marches. Ce paradigme permet également l'écriture aisée d'une borne sur les performances d'estimation du temps de propagation inter-capteur. Nous proposons dans la suite d'étudier ces performances, par la simulation, avec plusieurs configurations où seule la nature de  $p_\Delta$  change.

### 2.5.3 Illustrations du filtrage spatial des sources

Quatre configurations sont étudiées pour rendre compte des effets de différentes distributions réalistes de sources :

1. Dans la première configuration, nous observons le cas d'une distribution uniforme de différences de marche ; ce qui correspond au cas où l'identité de Ward a été retrouvée ;
2. La deuxième configuration reprend la première, mais avec un petit nombre de sources ;
3. La troisième configuration met en oeuvre des sources induisant des différences de marche tirées uniformément dans un sous-espace de  $[-1, +1]$ , correspondant par exemple à la restriction des angles d'incidence des front d'onde à un angle solide donné ;
4. La dernière configuration met en oeuvre des sources induisant des différences de marches qui ne sont pas tirées uniformément, ce qui correspond à privilégier certaines directions plutôt que d'autres.

Dans la simulation mise en oeuvre, les différences de marches sont comprises entre  $-1$  et  $+1$ , sachant que la distance inter-capteur est  $1$  m et que la vitesse de propagation  $c_0$  est fixée à  $1$  m/s. Le temps de propagation inter-capteur  $t_1$  est donc égal à  $1$  s.

Les sources sont modélisées par un spectre plat sur une bande  $B$  et nul en dehors, conduisant la fonction d'auto-corrélation de la source à prendre la forme d'un sinus cardinal. La bande passante  $B$  est choisie grande, égale à  $1$  kHz et la fréquence d'échantillonnage vaut  $5$  kHz. Comme suggéré dans l'algorithme d'identification passive, chaque corrélation est normalisée. La corrélation de Green est formée en sommant la corrélation des champs de chaque source. Comme l'expose l'Eq. (2.44), les corrélations des champs vus aux capteurs sont l'auto-corrélation de la source  $C_f$ , décalée dans le temps de la différence de marche  $\Delta/c_0$ . Le processus d'estimation de la corrélation de Green est simulé en ajoutant aux corrélations un bruit additif gaussien centré de variance  $0.01$ , correspondant à  $n(t)$  dans le modèle Eq. (2.45).

Dans les Fig. 2.21 – 2.24, les corrélations de Green estimées sont présentées, associées aux fonctions de Green estimées par application de l'identité de Ward. Les distributions suivies par les angles d'incidence sont précisées sous la forme d'un cercle trigonométrique, où l'intensité du cercle représente l'amplitude de la densité ; la barre de couleur précise cette carte d'intensité.

Notons  $N$  le nombre de sources employées et  $L$  la longueur  $d(\mathbf{x}, \mathbf{y})/c_0$ . Nous observons que tant que  $NL \gg 2t_1$ , dériver la corrélation de Green conduit à la présence d'un pic à 1 sec. La correspondance avec le modèle d'observation (2.13) est complète. Quand  $NL \leq 2t_1$ , la corrélation de Green n'est pas estimée correctement. En conséquence, la fonction de Green obtenue est très peu robuste, ne permettant pas d'estimer le temps de propagation de manière fiable. Un cas particulier du scénario Fig. 2.21 est celui où les angles d'incidence sont uniformément répartis dans un sous-espace de l'espace des phases. En pratique, c'est une configuration qui peut être rencontrée lorsqu'une interface entre deux milieux de propagation va contraindre la direction de propagation. C'est un cas qui a été observé par [20] dans un bassin California. Dans ce cas, l'information récupérée est filtrée spatialement de telle sorte que l'estimation de paramètres spatiaux s'en trouve affectée. En particulier, un biais peut apparaître comme c'est le cas dans la simulation :  $B(\hat{t}_1) = t_1/2$ , pour l'estimation du temps de propagation. Ce biais disparaît si le cas  $\theta = 0$  est contenu dans la distribution des sources, auquel cas l'estimation du temps de propagation peut être atteinte.

Il est important de noter à présent qu'un estimateur alternatif à l'approche qui exploite l'identité de Ward, consisterait, pour ce scénario de milieu infini, à mesurer le support temporel dans lequel les différences de marche se répartissent. Un estimateur du temps de propagation  $t_1$  pourrait alors être construit en ne retenant que la différence de marche la plus grande. Cet estimateur n'est pas étudié ici car il est construit sans faire appel à l'identité de Ward.

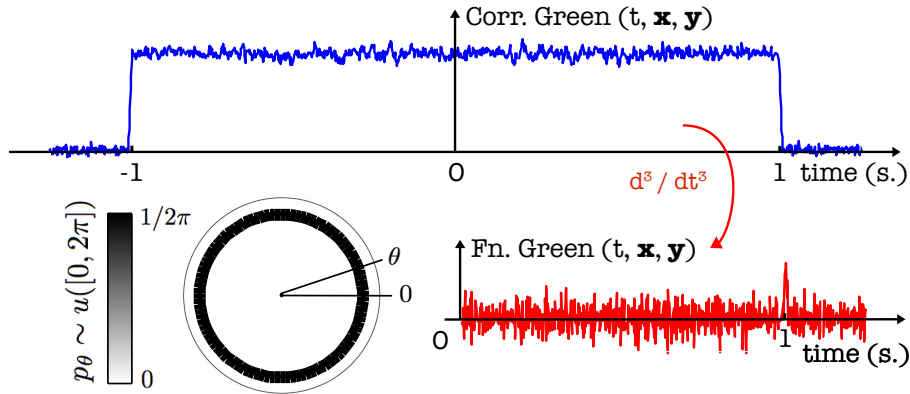


FIGURE 2.21 – Configuration 1. Scénario où les différences de marche sont tirées uniformément dans  $[-1; +1]$ . Nombre de sources simulées  $N = 10^5$ .

En conclusion, il apparaît que pour que l'estimateur de temps de propagation soit fiable et robuste, il est nécessaire qu'un grand nombre de sources ( $NL \gg 2t_1$ ) soit réparties uniformément dans l'espace des phases complet. Un écart à cette règle correspond à un filtrage spatial de sources blanches en espace (en vecteur d'onde) et résulte en un estimateur possiblement biaisé. Contrairement au filtrage temporel, qui résulte en une incertitude sur la position du temps de propagation, matérialisée par un résidu d'estimation (le plus souvent) centré, et de variance donnée, la blancheur spatiale du champ source est une condition *sine qua non* pour que l'estimateur du temps de propagation présenté ait du sens.

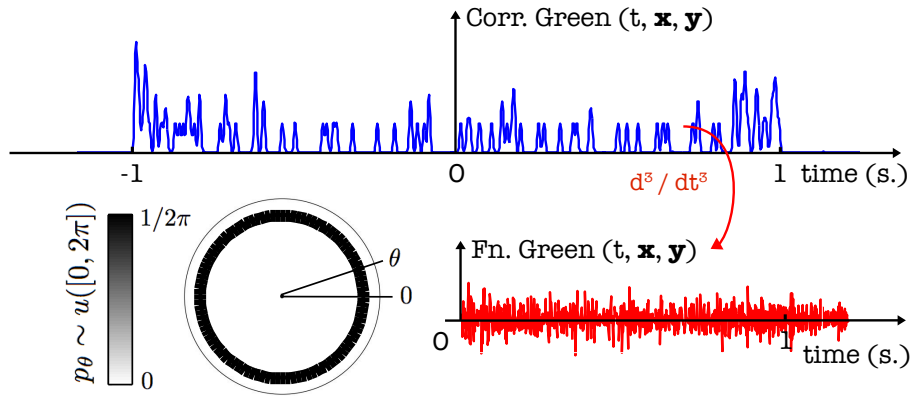


FIGURE 2.22 – Configuration 2. Scénario où les différences de marche sont tirées uniformément dans  $[-1; +1]$  mais pour un petit nombre de sources  $N = 10^2$ .

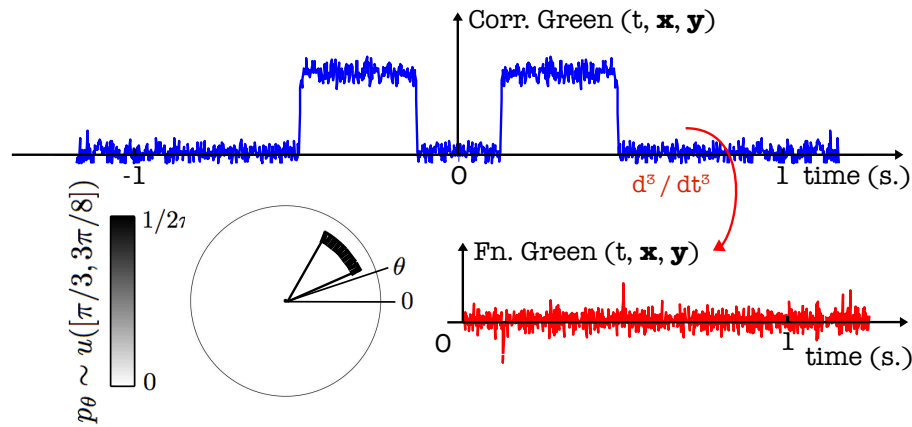


FIGURE 2.23 – Configuration 3. Scénario où les angles d'incidence sont contraints à un angle solide restreint  $[0.1; 0.3]$ .

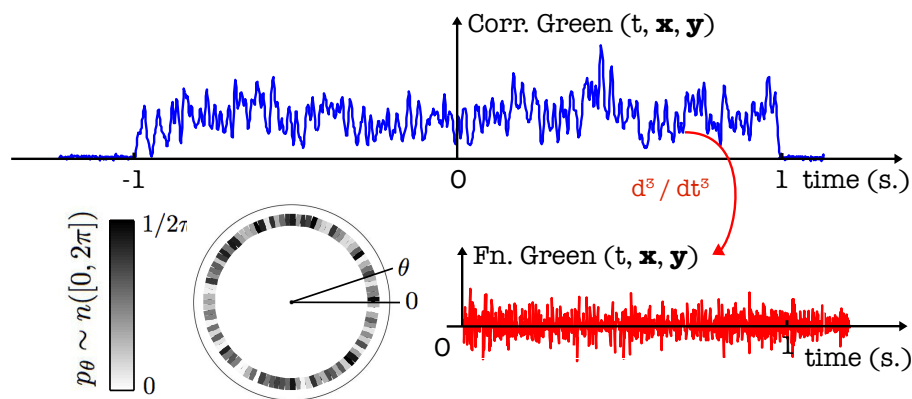


FIGURE 2.24 – Configuration 4. Scénario où les différences de marche ne sont pas tirées selon une loi uniforme. Nombre de sources :  $10^5$ .

### 2.5.4 Etude de performances conclusives

Dans cette section, deux modèles d'observation ont été proposés dans un contexte passif, où les champs utilisés ont été générés par des sources d'opportunité non contrôlées par l'utilisateur. Le premier modèle concerne la fonction de Green et s'exprime :

$$y(t) = \frac{c_0^2}{\alpha^2} C_f(t) * \delta(t - t_1) + w(t). \quad (2.51)$$

Le second modèle concerne la corrélation de Green, développé dans cette section et construit à partir d'un ensemble de sources d'ondes planes, spatialement paramétrées par leur angle d'incidence sur les capteurs :

$$z(t) = \int_{-1}^1 p_{\Delta}(u) C_f(t - u) du + n(t), \quad (2.52)$$

Il a été montré que ces deux modèles, bien que développés séparément, sont bel et bien reliés par l'identité de Ward (en considérant une dissipation constante). Le modèle sur la fonction de Green est un simple pulse à bande limitée, applicable autant dans un contexte d'estimation passive qu'active. La différence entre les deux approches se manifeste par une dégradation nette du RSB dans le cas passif, principalement attribuée aux calculs numériques qui correspondent à l'application de l'identité de Ward (en particulier la dérivation).

De ces deux modèles, deux bornes inférieures sur la variance de l'estimateur du temps de propagation  $t_1$  ont été proposées, voir Eq. (2.50) et Eq. (2.19). Ces bornes sont à présent comparées. La configuration pour le contexte l'estimation passive est choisie de telle sorte que la distribution spatiale n'induisse pas d'erreur de détection du temps de propagation  $t_1$  pathologiques ; c'est la configuration 1, voir Fig. 2.21 qui est appliquée, où les différences de marches sont tirées uniformément dans  $[-1; +1]$  et un très grand nombre de source est utilisé. Dans le cas actif, c'est le modèle Eq. (2.18) qui est employé, pour des sources de bande-passante plus ou moins grandes. L'écart-type moyen sur le temps de propagation  $t_1$ , estimé dans les contextes actif et passif est représenté Fig. 2.25.

La borne de Cramer-Rao théorique dans le contexte passif est deux décades plus bas que celle obtenue en simulation. Cette différence était prévisible, puisque la borne théorique a été dérivée à partir du modèle d'observation Eq. (2.45), qui décrit la corrélation de Green. En pratique, le temps de propagation n'a pas été estimé depuis la corrélation de Green puisque l'accès à cette information depuis la corrélation de Green n'est pas modélisé (si l'estimateur de support temporel est écarté). A la place, pour estimer le temps de propagation, il s'agissait d'abord d'estimer la fonction de Green, en appliquant l'identité de Ward Eq. (2.9), pour ensuite localiser le maximum. Ainsi, des baisses de performances de plusieurs décades étaient prévues, dues entre autre à la dérivation triple et à l'extraction de l'argmaximum.

Les performances dans le contexte actif d'estimation sont sensiblement meilleures que dans le contexte passif, mais chacune rejoint la borne expérimentale déterminée par la résolution d'échantillonnage et le nombre d'itérations utilisées pour calculer les statistiques. Il est notable que ces performances, pour être améliorées, bénéficieraient grandement d'une technique

**Légende** (Figure 2.25)

*Ecart-type de l'estimateur du temps de propagation. (lignes en gras) (a) Performances de l'estimateur passif. (a') Borne théorique dans le cas passif, dérivée du modèle d'observation Eq. (2.50). (pointillés) (b) Performances de l'estimateur actif, à partir du modèle (2.18). (b') Borne théorique du modèle de la fonction de Green (2.19), applicable pour l'estimateur passif et actif.*

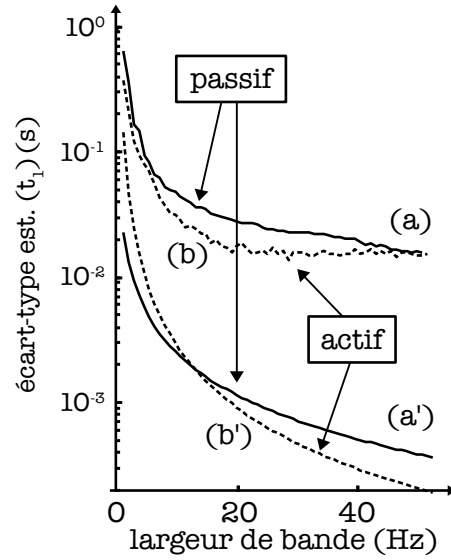


FIGURE 2.25 – Comparaison des performances.

d'estimation d'argmax super-résolution. De telles techniques permettraient, à moindre coût en terme de complexité algorithmique pour une simple interpolation par exemple, d'aller au-delà de la notion de résolution d'échantillonnage dans le calcul des statistiques de l'estimateur.



# Références

- [1] G. PENELET, “Acoustic en fluide visqueux et conducteur de la chaleur”, rap. tech., 2014 (cf. p. 23).
- [2] G. G. STOKES, “G.G. Stokes, Trans. Cambridge Philos. Soc. 8 287 (1845).”, *Transactions Cambridge Philos. Soc.*, t. 8, n° 287, 1845 (cf. p. 23).
- [3] J.-L. LACOUME, “Tomographie passive : Observer avec du bruit”, dans *GRETSI*, 2007, p. 601–604 (cf. p. 25, 54).
- [4] R. WEAVER et O. LOBKIS, “On the emergence of the Green’s function in the correlations of a diffuse field : pulse-echo using thermal phonons”, *Ultrasonics*, t. 40, n° 1-8, p. 435–439, mai 2002 (cf. p. 25).
- [5] P. GOUÉDARD, L. STEHLY, F. BRENGUIER, M. CAMPILLO, Y. C. D. VERDI, E. LAROSE, L. MARGERIN, P. ROUX et R. L. WEAVER, “Cross-correlation of random fields : mathematical approach and applications”, *Geophysical Prospecting*, n° 56, p. 375–393, 2008 (cf. p. 25).
- [6] Y. COLIN DE VERDIÈRE, “Semiclassical analysis and passive imaging”, *Nonlinearity*, t. 22, n° 6, R45–R75, juin 2009 (cf. p. 25).
- [7] M. CARMONA, O. J. J. MICHEL, N. SPRYNSKI, J.-L. LACOUME et B. NICOLAS, “Ward identities for visco-acoustic and visco-elastic propagation media”, *Wave Motion*, t. 49, n° 4, p. 484–489, 2012 (cf. p. 25).
- [8] A. SCHUSTER, *Theory of optics*, EDWARD ARNOLD, éd. Longon, 1909 (cf. p. 27, 34).
- [9] H. L. VAN TREES, K. L. BELL et Z. TIAN, *Detection, estimation and filtering theory*, WILEY, éd. 1968 (cf. p. 29, 30, 54).
- [10] S. M. KAY, *Fundamentals of statistical signal processing*. Massachussets, 2013 (cf. p. 30).
- [11] J. ZIV et M. ZAKAI, “Some Lower Bounds on Signal Parameter Estimation”, *IEEE Transactions on Information Theory*, t. 15, n° 3, p. 386–391, 1969 (cf. p. 30, 31).
- [12] K. AKI et B. CHOUET, “Origin of Coda Waves : source, attenuation, and scattering effects”, *Journal of Geophysical Research*, t. 80, n° 23, p. 3322–3342, 1975 (cf. p. 32).
- [13] M. CAMPILLO et A. PAUL, “Long-Range Correlations in the Diffuse Seismic Coda”, *Nature Reports*, t. 299, n° January, p. 547–549, 2003 (cf. p. 32).
- [14] O. I. LOBKIS et R. L. WEAVER, “On the emergence of the Green’s function in the correlations of a diffuse field”, *Acoustics 40*, t. 110, n° September, p. 3011–3017, 2001 (cf. p. 32).
- [15] G. BAL et L. RYZHIK, “Time splitting for wave equations in random media”, *ESAIM : Mathematical Modelling and Numerical Analysis*, t. 38, n° 6, p. 961–987, 2004 (cf. p. 33).
- [16] R. SNIEDER, K. WAPENAAR et U. WEGLER, “Unified Green’s function retrieval by cross-correlation ; connection with energy principles”, *Physical Review E*, t. 75, n° 3, p. 036 103, mar. 2007 (cf. p. 33).



- [17] S. M. HANASOGE, “The influence of noise sources on cross-correlation amplitudes”, *Geophysical Journal International*, t. 192, n° 1, p. 295–309, 2013. arXiv :arXiv:1210.2313v1 (cf. p. 33).
- [18] K. WAPENAAR, D. DRAGANOV, R. SNIEDER, X. CAMPMAN et A. VERDEL, “Tutorial on seismic interferometry : Part 1 — Basic principles and applications”, *Geophysics*, t. 75, n° 5, 75A195–75A209, sept. 2010 (cf. p. 33).
- [19] M. FINK, “Time Reversal of Ultrasonic Fields”, *IEEE Transactions on Ultrasonics*, t. 39, n° 5, 1992 (cf. p. 33).
- [20] G. a. PRIETO, J. F. LAWRENCE et G. C. BEROZA, “Anelastic Earth structure from the coherency of the ambient seismic field”, *Journal of Geophysical Research*, t. 114, n° B7, B07303, juil. 2009 (cf. p. 33, 56).
- [21] BLESSER, “Working Vocabulary for the Experience of Aural”, rap. tech. January, 2007, p. 1–6 (cf. p. 33).
- [22] G. DEFRANCE et J. POLACK, “Measuring the mixing time in auditoria”, dans *Acoustics*, Paris, 2008, p. 3871–3876 (cf. p. 34, 37).
- [23] O. LEGRAND et F. MORTESSAGNE, *Wave chaos for the Helmholtz equation*, Cambridge. 2010, p. 22–41 (cf. p. 35).
- [24] C. FRÉDÉRIC, “Peut-on entendre la forme d’un tambour?”, rap. tech. (cf. p. 44).
- [25] M. CARMONA, “Identification passive des milieux. Application à la reconstruction de géométries dans les réseaux de capteurs et au diagnostic des structures”, thèse de doct., Université de Grenoble, 2008 (cf. p. 44).
- [26] C. R. FARRAR et G. H. JAMES, “System identification from ambient vibration measurements on a bridge”, *Journal of Sound and Vibration*, t. 205, n° 1, p. 1–18, 1997 (cf. p. 45).

# Applications au SHM

Dans ce chapitre, nous appliquons le savoir-faire développé autour de l'identification passive, dans le domaine du SHM (Structural Health Monitoring). Nous montrerons que la modalité acoustique permet d'atteindre les mêmes objectifs que la modalité sismique.

		sismique		acoustique	ultrasons
		0 Hz	20 Hz	20 kHz	
		accéléromètres		microphones	
techniques	déformée	✓ données tangentielles Ex : Sprynski et al. [a]	✓ localisation d'un réseau Ex : Vincent et al. [c]	✗	
	vibratoire	✓ output-only modal id. Ex : Farrar et al. [b]	✓ corrélation de codas Ex : Vincent et al. [d]	✗	
	dét. défauts	✗	✗	✓	Ex : Larose et al. [e]

[a] - Surfaces reconstruction via inertial sensors for monitoring, *EWSHM 2014*.

[b] - System identification from ambient vibration measurements on a bridge, *J. of Sound & Vibration 1, 1999*.

[c] - Passive acoustic sensor network localization, app. to structural geometry monitoring, *EWSHM 2014*.

[d] - SHM using acoustic sources of opportunity : monitoring Green's correlations, *IWSHM 2015*.

[e] - Locating a Small Change in a Multiple Scattering Environment, *Applied Physics Letters 96, 2010*.

La thématique SHM est à enjeux technologiques, et celui de l'autonomie d'un noeud de mesure est très fort. Dans cette optique, nous travaillerons sur la réduction de la quantité de données manipulée. Pour ce faire, nous tirerons profit du fait que des fonctions de corrélations peuvent être estimées à partir de champs gaussiens fortement quantifiés. Ce faisant, nous rendrons compatibles les notions de noeud de mesure autonome et de SHM passif acoustique et c'est ainsi, en guise d'ouverture à ce chapitre, qu'une description fonctionnelle sera faite de ce que serait un tel noeud de mesure, baptisé noeud A. P. I. (Acoustic Passive Identifier).

## Sommaire

<b>3.1</b>	<b>Estimation à partir de signaux quantifiés</b>	<b>65</b>
<b>3.2</b>	<b>Applications de la modalité acoustique pour SHM</b>	<b>66</b>
3.2.1	Cas d'une poutre avec N capteurs	66
3.2.2	Cas d'un milieu confiné avec 1 capteur	72
<b>3.3</b>	<b>Ouvertures vers un identificateur passif</b>	<b>76</b>
3.3.1	Vers un noeud de mesure API	76
3.3.2	Toujours plus loin pour l'autonomie	78



### 3.1 Estimation à partir de signaux quantifiés

Dans la course à l'autonomie d'un système, le principal objectif est la réduction des données à acquérir, traiter et transmettre. Nous proposons de travailler sur tous ces pans à la fois, en considérant la problématique d'estimation de fonctions de corrélation sur des signaux très faiblement quantifiés. La réduction proposée est même drastique, puisqu'il s'agira de ne garder que le signe du champ mesuré et ainsi de ne manipuler que des signaux quantifiés sur 1 bit. La réduction typique induite sur la quantité de données stockées et traitées est d'un facteur 16 si chaque échantillon était stocké sur 2 octets à l'origine. A l'acquisition, ce choix annule la nécessité d'utiliser un convertisseur analogique classique (sur 12 bits exemple), alors qu'un simple comparateur peut faire l'affaire. La consommation globale du noeud en sera *de facto* améliorée.

Réduire la quantification des champs acquis était une problématique liée aux capacités de traitement numériques limitées de l'époque, et soixante ans plus tard, les procédés employés ressurgissent afin de permettre l'intégration de modalités « data-phage » comme l'acoustique audible. Tout particulièrement, notre protocole d'identification à partir de sources d'opportunité repose sur la corrélation de codas acoustiques. La corrélation de signaux quantifiés peut être perçue comme une version distordue de la corrélation vraie et lorsque la quantification est faible, les efforts dans ce domaine de la recherche consistaient à formaliser et à étudier cette perturbation, voire à lui trouver une compensation sous certaines hypothèses.

Les travaux de Van Vleck [1] assoient des bases théoriques de cette thématique en généralisant l'approche à la quantification sur  $N$  bits d'une fonction de corrélation donnée, et vient ainsi résumer et compléter des travaux plus anciens comme ceux de Wolff [2], faisant état de fait des performances des corrélateurs dits « à coïncidence de phase ». Suite à ceci, une vague des travaux vient compléter l'étude des statistiques des estimateurs de corrélations de champs quantifiés, voir [3, 4, 5]. Par ailleurs, notons que des techniques comme le dithering (bruit rajouté à la quantification pour simuler une quantification plus grande qu'elle n'est réellement), classiques en traitement du signal acoustique, sont plus tard éprouvées dans le contexte de la reconstruction de fonctions de corrélations ; voir par exemple [6]. Dans le cadre de nos travaux, la coda acoustique est modélisée par une variable gaussienne centrée, dont la fonction de corrélation est liée à la corrélation de Green par une relation de filtrage. Rappelons ici que cette relation dépend entre autre de la chaîne d'acquisition, de la source impulsionnelle (son contenu spectral) et de l'état de réverbération qui réduit la largeur de bande utile de la coda par dissipation et atténuation géométrique. Sous cette hypothèse de gaussianité, il existe un lien fonctionnel entre la corrélation de la coda quantifiée et de sa corrélation vraie. Soient  $x(t)$  et  $y(t)$  les codas à corréler, supposées centrées, alors :

$$\frac{\mathbb{E}[x(t)y(t-\tau)]}{\sqrt{\mathbb{E}[x^2(t)]\mathbb{E}[y^2(t)]}} = \sin\left(\frac{\pi}{2}\mathbb{E}[\text{signe}(x(t))\text{signe}(y(t-\tau))]\right). \quad (3.1)$$

La preuve de cette relation peut être trouvée dans [1]. Une première remarque est que cette transformation conserve la position du maximum d'amplitude, ce qui pourra s'avérer crucial par la suite lorsque nous chercherons à estimer des distances inter-capteurs. La deuxième

remarque est que l'énergie du champ est normalisée et que cette information est donc perdue. Rappelons cependant que la normalisation est dans notre protocole une étape employée pour égaliser les contributions de coda, différentes suivant leur état de dissipation plus ou moins avancé. Enfin, notons que la correction  $y = \sin(x)$  est une relation peu complexe et donc globalement adaptée au calcul embarqué.

Par la suite, le protocole d'estimation passive sera conduit en calculant des corrélations de codas acoustiques quantifiées sur 1 bit et corrigées par l'application de l'Eq. (3.1).

## 3.2 Applications de la modalité acoustique pour SHM

Le protocole d'identification passive développé jusqu'à maintenant sera mis en oeuvre dans des scénarios réalistes du monde du Structural Health Monitoring. Nous montrerons que la modalité acoustique permet d'atteindre les mêmes objectifs que la modalité sismique : analyse vibratoire (estimation modale) et reconstruction de déformée (par exemple, le profil en 2 ou 3 dimensions d'une interface entre deux milieux de propagation, tels que l'air et un acier). Les fonctionnalités présentées devront être réalisables avec le noeud de mesure API présenté dans la section 3.3.1 et ce faisant, son utilité, sa légitimité et son intérêt auront été justifiés.

Deux scénarios sont proposés :

1. Le suivi des variations du module d'Young d'une poutre, élément structurel fondamental dans l'ingénierie des structures ;
2. Le monitoring d'un milieu clos de type « bureau open-space ».

Les fonctionnalités présentées dans ce chapitre exploitent un nombre variable de capteurs A.P.I., qui seraient amenés à faire appel à l'ensemble de leur fonctions : calcul embarqué de corrélation de champ, de fonction de Green et/ou transfert complet des champs mesurés. Les contraintes technologiques du noeud imposent que les performances présentées soient évaluées avec 10 ko de données pour chaque noeud, soit une demi-seconde d'enregistrement à 44.1 kHz, en utilisant la technique précédemment introduites de calcul de corrélations sur des signaux quantifiés sur 1 bit.

### 3.2.1 Cas d'une poutre avec N capteurs

Dans la première expérience, le milieu étudié est une poutre, mesure 3 mètres de long et est posée sur deux appuis éloignés de 1 mètres des extrémités. La poutre est en acier et possède un module d'Young et un moment quadratique, au repos, qui valent respectivement  $E_0 = 210 \times 10^6 \text{ N/m}$  et  $I = 3.9 \times 10^{-8} \text{ m}^4$ . Un jeu de 7 capteurs  $C_i$ ,  $i = 1, \dots, 7$  est déployé sur la poutre. Dans la Fig. 3.1, la configuration de l'expérience est représentée. Notons que les microphones enregistrent le champ acoustique qui se propage dans la pièce, et non dans la poutre. Le couplage entre les deux milieux existe mais est négligé pour cette étude.

Dans la matrice Eq. (3.2) sont représentés les résultats d'estimation des distances inter-

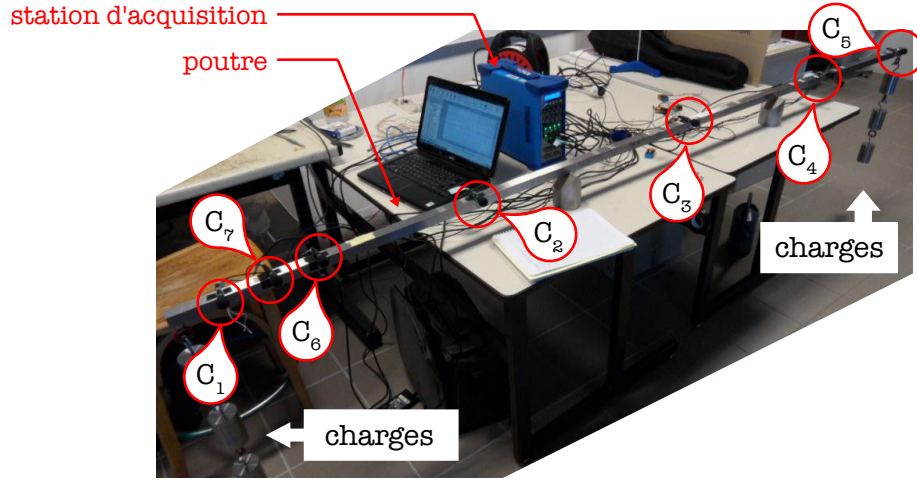


FIGURE 3.1 – Configuration expérimentale du monitoring d'une poutre.

capteurs. Le protocole employé pour parvenir à ces estimations est l'application de l'identité de Ward visco-acoustique sur le signe des codas issues d'événements d'opportunité, puis extraction du temps de propagation sur la fonction de Green estimée. Les valeurs vraies sont en noir (en cm) et les erreurs d'une réalisation de chaque estimation figure en rouge :

$$\begin{matrix} & C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 \\ \begin{matrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{matrix} & \left( \begin{array}{ccccccc} 0 & 60+0.4 & 140-0.1 & 220-0.1 & \times & 20+1.3 & 10+0.6 \\ & 0 & 80+1.6 & 160+0.2 & 220+0.5 & 40-0.6 & 50-0.3 \\ & & 0 & 80+0.3 & 140+0.6 & 120+0.1 & 130+0.7 \\ & & & 0 & 60+0.7 & 200+0.4 & 210-1.1 \\ & & & & 0 & 260-0.7 & 270-1.4 \\ & & & & & 0 & 10-0.1 \end{array} \right) \end{matrix} \quad (3.2)$$

L'estimateur de distances inter-capteurs dans un contexte passif est non biaisé et est accompagné, malgré la restriction forte sur la quantité de données utilisée, d'une erreur de l'ordre de la résolution spatiale et également de la taille des capteurs. L'écart-type global, toutes distances confondues, est en effet de 0.7 cm. En pratique, l'estimation des distances inter-capteurs s'est montrée robuste sur plusieurs mètres, faute de pouvoir instrumenter sur de plus grandes distances.

A partir de ce jeu de distances, le monitoring peut être réalisé de plusieurs manières. La première consiste à estimer le module d'Young de la poutre et de le suivre lorsque cette dernière change de géométrie, voir par exemple [7]. En effet, à partir du modèle de poutre reposée sur 2 appuis, la mesure d'une seule distance inter-capteurs permet d'estimer le module de la poutre. Ainsi, les variations de cette distance permettront d'estimer les variations du module d'Young. Ici, nous proposons d'estimer le module d'Young en exploitant les capteurs positionnés aux extrémités de la poutre, puisque ce sont ces endroits de la poutre qui se déplacent le plus lorsqu'elle est chargée. L'approche présentée reste cependant valable pour n'importe quelle paire de capteurs choisie.

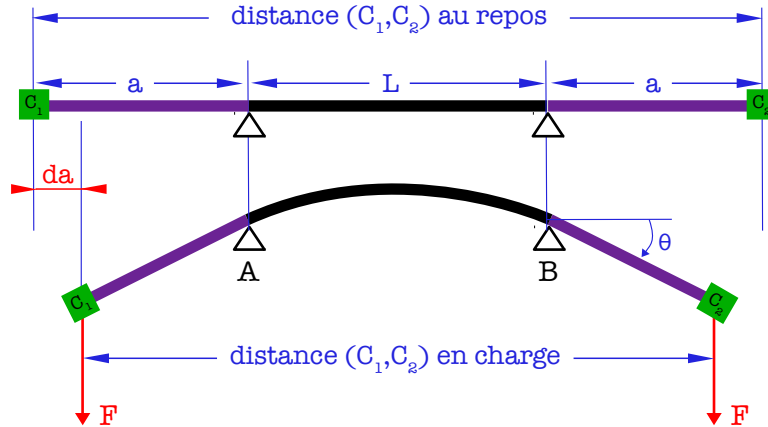


FIGURE 3.2 – Schéma de la poutre en position repos et chargée.

Expérimentalement, le changement de géométrie de la poutre est obtenue en la chargeant à ses extrémités. Une fois chargée, la distance aérienne qui sépare les capteurs diminue. Dans la configuration étudiée, voir Fig. 3.2, l'ajout d'une charge  $\Delta F = F - F_0$  ( $F_0$  charge à vide) identique aux extrémités induit une variation du moment de flexion constant entre les appuis, égal à  $-\Delta Fa$ . Si l'on dote le schéma d'un repère plan  $(\vec{x}, \vec{y})$  orthonormé direct où  $\vec{x} = \vec{AB}$ , alors entre les points  $\vec{A}$  et  $\vec{B}$ , la poutre est décrite par l'équation :

$$y(x) = \frac{-\Delta Fa}{2EI} \times (x^2 - Lx), \text{ avec } y(0) = y(L) = 0 \text{ par exemple,} \quad (3.3)$$

alors, la flèche de la poutre au centre de  $[A, B]$  vaut  $\Delta FaL^2/8E_0I$ , ce qui équivaut à la présence d'une force en ce point, avec une norme égale à  $6\Delta Fa/L$ . L'angle  $\theta$  induit par cette force, aux points d'appui, vaut  $3FaL/8E_0I$ . Finalement en supposant que les segments en dehors des points d'appui se déforment peu, l'approximation suivante est vérifiée :

$$\Delta d(C_1, C_2) = 2da \approx 2a (1 - \cos(\theta)) = 2a \left( 1 - \cos \left( \frac{3aL}{8E_0I} \Delta F \right) \right). \quad (3.4)$$

Ainsi formulée, nous pouvons faire évoluer la relation qui relie la variation de la distance inter-capteurs avec la variation de la charge, pour faire correspondre à la place une variation du module d'Young, de manière à retrouver une poutre non chargée. Il s'agit simplement d'écrire que le nouveau module d'Young  $E$  vaut  $F_0/F - F_0$ . Par ce changement de variable, les variations du module d'Young sont maintenant directement observables à travers les variations de la distance inter-capteurs :

$$\Delta E = \frac{8I}{3aLF_0} \arccos \left( 1 - \frac{\Delta d(C_1, C_2)}{2a} \right). \quad (3.5)$$

Dans la Fig. 3.3, cette relation est représentée de manière à faire apparaître les seuils où la variation de la distance inter-capteurs est détectable par la résolution spatiale. En correspondance figure la variation de module d'Young. A 44.1 kHz, la résolution spatiale est de 7.8

mm puisque l'onde se propage à 343 m/s, ce qui correspond également à la taille des capteurs. Pour que les capteurs se rapprochent de cette distance, la poutre devrait voir son module d'Young passer en dessous de  $0.39 E_0$  en étant chargée à 10kg. C'est le seuil de détectabilité d'une variation du module d'Young. Il correspond à l'apparition d'une flèche de 4.3 mm. Une fois ce seuil dépassé par la variation de la distance inter-capteur, chaque multiple de ce seuil donne accès à une nouvelle mesure de la variation du module d'Young.

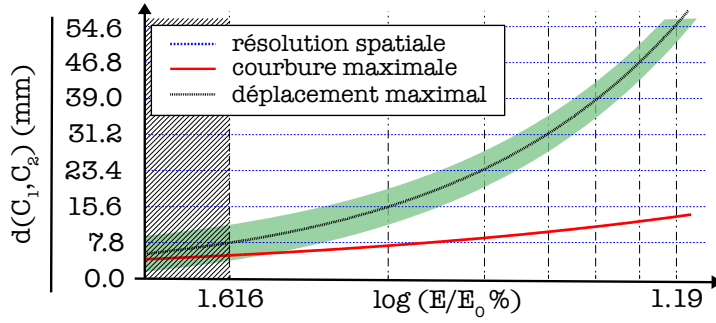


FIGURE 3.3 – bending and detection

**Légende (Figure 3.3)**

*Variation de la distance inter-capteurs en fonction de la variation relative du module d'Young. La zone grisée est en dessous du seuil de détection dû à la résolution spatiale. Les paliers de détections correspondent à des variations de la flèche qui sont millimétrique.*

Le seuil de détectabilité du module d'Young paraît grand, puisqu'il représente une perte de 60% de sa valeur. Il faut cependant garder en tête que le modèle est à l'échelle d'une salle de laboratoire, où les dimensions de la poutre sont une réduction de celles des poutres usuellement rencontrées dans l'ingénierie des structures, comme un pont par exemple. Quant à elle, la résolution spatiale ne varie pas puisqu'elle ne dépend pas de la géométrie du milieu. Ainsi, le seuil de détectabilité des variations du module d'Young sur des géométries plus grande sera amené à diminuer considérablement.

En conclusion de cette première approche, il apparaît que l'exploitation de deux capteurs permet de suivre l'évolution du module d'Young, moyennant la modélisation de la poutre et la connaissance de la localisation des capteurs. La précision des estimations, malgré la quantité restreinte de données, offre une confiance maximale puisque de l'ordre de la taille des capteurs. Cette précision, qui ne dépend pas de la structure, offre un suivi des paramètres structuraux que nous pouvons supposer d'autant plus précis relativement que la structure est grande. Ainsi, pour des poutres caractérisant un pont par exemple, atteignant la dizaine de mètres de longueur, il n'est pas aberrant de pouvoir espérer détecter des variations du module d'Young de l'ordre du pour-cent.

**3.2.1.1 Des distances aux positions**

L'approche présentée dans la suite est destinée à exploiter un ensemble de distances inter-capteurs, afin d'accéder à la géométrie du milieu. En terme de complexité algorithmique, c'est l'algorithme d'estimation sujet à une propagation d'erreurs la plus forte présentée dans la totalité de ce manuscrit. En effet, estimer la géométrie du réseau de capteurs sera présentée sous forme d'un problème inverse traitant un jeu de distances inter-capteurs. Ces distances auront été obtenues par l'algorithme d'estimation dans un contexte passif acoustique, exploitant la



coda des évènements impulsionnels acoustiques. Pour estimer ces distances, il aura fallu estimer des corrélations de Green, puis appliquer l'identité de Ward, puis enfin extraire l'argmaximum. Ainsi, les erreurs d'estimation associées, caractérisées dans le chapitre précédent du manuscrit, seront à nouveau propagées dans un algorithme qui va résoudre un problème inverse. Dans ce genre de problème, il est évident que le comportement des incertitudes aux entrées de l'algorithme est crucial pour déterminer celles en sortie.

Reconstruire la géométrie d'un réseau de capteurs peut avoir plusieurs significations et pour nous, il s'agira d'estimer son profil dans un repère  $(0, \vec{x}, \vec{y}, \vec{z})$ , à partir des positions  $\mathbf{X}$  estimées des capteurs. En manipulant avec un réseau à  $N$  capteurs, l'algorithme d'estimation dans un contexte passif permet d'estimer un nombre  $M$  de distances inter-capteurs parmi les  $(N-1)(N-2)/2$  totales. A la vue des résultats présentés jusqu'ici, ces estimations sont non biaisées et sont perturbées par un bruit d'estimation additif modélisé gaussien de variance  $\sigma^2 = 0.49 \text{ cm}^2$ . Les distances estimées sont concaténées dans une matrice  $\hat{D}$  :

$$\hat{D} = \mathbf{A} \otimes (\mathbf{D}(\mathbf{X}) + \mathbf{B}) \quad (3.6)$$

où :

- $\mathbf{D}(\mathbf{X})$  est la matrice des véritables distances inter-capteurs, qui dépend de leurs positions  $\mathbf{X}$ . Typiquement,  $\mathbf{D}_{i,j} = d(\mathbf{X}_i, \mathbf{X}_j) = \|\mathbf{X}_i\|^2 + \|\mathbf{X}_j\|^2 - 2\langle \mathbf{X}_i, \mathbf{X}_j \rangle$  ;
- $\mathbf{A}$  est la matrice d'adjacence de la matrice, dont les valeurs binaires déterminent si une distances a pu être estimée ou non.  $\otimes$  est le produit de Hadamar ;
- $\mathbf{B}$  est une matrice symétrique à diagonale nulle où chaque élément est le tirage d'une variable aléatoire centrée de même variance  $\sigma^2$ , traduisant la confiance associée à une mesure de distance inter-capteur, choisie identique ici pour toute les paires de capteurs.

L'ordre des choses pour estimer la géométrie d'un réseau de capteurs est premièrement de discriminer les capteurs localisables de ceux qui ne le sont pas. Alors seulement se posera le choix de l'algorithme qui s'appuiera sur les distances liant les noeuds localisables, afin d'estimer les positions des capteurs. Finalement, un algorithme de régression utilisera ces positions pour suivre la déformée du réseau.

Tout d'abord, la localisabilité d'un noeud de mesure est conditionnelle à un repère et ce repère peut être construit par d'autres noeuds du réseau. C'est l'angle d'approche suivi ici. En pratique, la nécessité d'une ancre absolue qui permette de localiser un réseau de capteurs au sein d'une structure dépend de l'application. Il est intéressant de noter qu'observer une variation de la géométrie du réseau ne requiert pas la localisation absolue du réseau mais rend néanmoins compte de l'évolution de la santé de la structure. Un capteur est dit localisable si l'ensemble de ses positions possibles est fini dénombrable. Un tel capteur est appelé ancre et un capteur localisable par un ensemble d'au moins  $d$  ancres, dans un espace (vectoriel) à  $(d-1)$  dimensions, est lui-même une ancre. Un réseau d'ancres est ainsi un ensemble de capteurs auto-localisables.

Le problème de localisation à partir d'une matrice de distances, consiste à déterminer  $\mathbf{X}$  sachant  $\mathbf{D}$ . Nous citerons ici que la résolution analytique du cas complet  $M = (N-1)(N-2)/2$  a été proposée par Torgerson [8, 9], qui traite en fait une classe de problème plus large dans le domaine de l'analyse de données : le multi-dimensional scaling. C'est Kruskal [10] qui introduit

la prise en compte des incertitudes sur les mesures de distances. Un résultat central à l'approche est la caractérisation de la stabilité spectrale d'une matrice perturbée, par exemple décrite dans le théorème de Davis-Kahan, voir [11]. Afin de traiter le cas non complet, c'est-à-dire  $M \leq (N - 1)(N - 2)/2$ , nous capitalisons sur les travaux de thèse de Mikael Carmona [12], dans lesquels il développe un algorithme itératif qui d'ailleurs a donné lieu à un brevet, voir [13]. L'algorithme est un algorithme glouton qui va, de manière récursive, ajouter à ses ancres tous les noeuds qui sont devenus localisables grâce aux dernières ancres ajoutées. Ainsi, pour un jeu d'ancres initial données, le réseau d'ancres le plus grand possible est trouvé. Pour déterminer si un ensemble de  $N$  capteurs forme un réseau d'ancres complet, il suffit de trouver au moins une configuration initiale qui conduisent à la localisabilité de l'ensemble des capteurs. L'application de l'algorithme de Kruskal permet alors d'estimer les positions des capteurs.

La Fig. 3.4 caractérise la localisabilité d'un réseau de  $N$  capteurs en fonction du nombre de distances inter-capteurs disponibles. Les statistiques sont calculées en simulant plusieurs réseaux de capteurs de taille variable, où la probabilité qu'une distance soit disponible est une variable de Bernoulli d'espérance contrôlée, identique pour chaque paire de capteur. Cette caractérisation permet d'identifier que plus le nombre de noeuds est grand, moins la densité de connections doit être forte pour que la probabilité que le réseau soit localisable. C'est un résultat intuitif dans le sens où un noeud est localisable dès lors qu'un nombre de distances supérieur à  $d + 1$  (dans un espace vectoriel de dimension  $d$ ) le concernant est disponible. Cette limite est une constante et est donc plus facilement atteinte lorsque le nombre de noeuds du réseau augmente. Ici, le rayon de visibilité d'un noeud n'est pas une composante prise en compte, et ainsi les noeuds voient à l'infini.

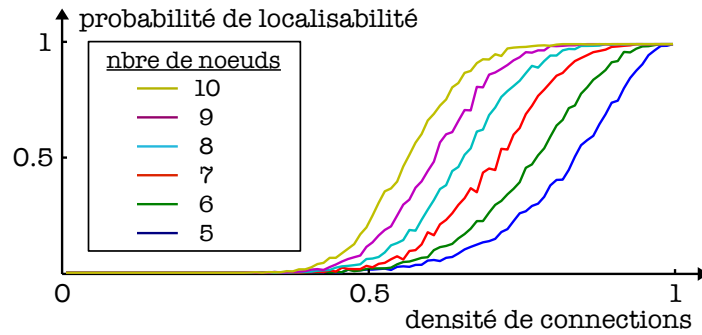


FIGURE 3.4 – Localisabilité du réseau en fonction de la densité de connections.

Pour déterminer la confiance à apporter aux positions estimées par l'algorithme de type *Multi Dimensional Scaling*, la propagation d'erreur dans l'algorithme de Kruskal a été caractérisée par Mikael Carmona dans ses travaux de thèse [12]. Dans l'algorithme, les contributions référencées à [14, 15] permettent d'identifier l'isométrie qui relie les deux réseaux formés par les ancres. C'est son application qui permet de faire coïncider le réseau de référence avec le réseau perturbé reconstruit.

La dernière étape consiste, à partir des positions des capteurs, à déterminer la courbe (2d ou 3d) qui régresse ou interpole le mieux ces positions, selon un critère à définir. Le seul *a priori* sur ce profil est qu'il est au moins  $\mathcal{C}_1$ . Une interpolation par splines cubiques par

morceaux est classiquement employée dans ce contexte là, mais ne fait pas intervenir de la notion de propagation d'erreur. A la place, un problème d'optimisation de type régression, voir par exemple [16], donc contraint par la confiance attribuée aux positions semble plus en adéquation avec le problème.

Le résultat (2d) de la reconstruction de la géométrie de la poutre à partir des distances Eq.(3.2) est présenté dans la figure Fig. 3.5. Les données ont été calibrées pour que la position des capteurs centraux soit identiques. La superposition des courbes ne requiert alors plus que l'application d'une isométrie, déjà introduite précédemment.

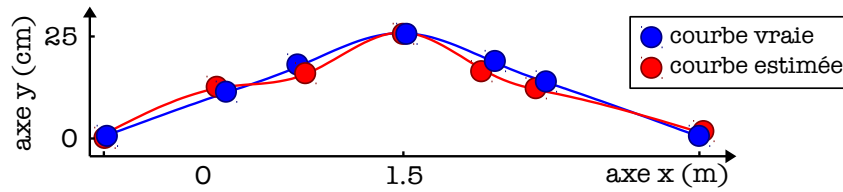


FIGURE 3.5 – Exemple de reconstruction de la géométrie de la poutre

Ici, étant données les dimensions de la poutre, l'erreur sur l'estimation de la géométrie n'est pas acceptable. Ce résultat peut cependant être modéré en considérant l'augmentation de la dimension caractéristique du milieu. En effet, la confiance sur l'estimation des distances ne dépend que très peu de la distance elle-même. Ainsi, l'erreur relative diminue linéairement avec l'augmentation de la distance. C'est cette erreur relative qui joue un rôle dans l'algorithme d'inversion et plus elle sera petite, plus la reconstruction sera fidèle. Par ailleurs, l'augmentation de la quantité de données utilisées. Ici, les conditions d'utilisation du noeud API limitent fortement la quantité de données utilisées. Il n'est exclu que la convergence des estimateurs puisse être améliorée.

### 3.2.2 Cas d'un milieu confiné avec 1 capteur

Il a été souligné que les fonctionnelles estimées à partir des codas acoustiques sont des versions filtrées des fonctionnelles de Green et leur estimation a été prouvée robuste. C'est cette propriété que nous exploitons à présent, à nouveau avec le milieu cuboïde Fig. 2.12. Un seul microphone est utilisé et trois configurations sont étudiées :

- Dans la configuration originale, le capteur est à une position de référence ;
- Dans le premier cas d'étude, la géométrie du milieu est changée, par l'ajout d'une boîte de dimensions  $10\text{cm} \times 10\text{cm} \times 10\text{cm}$ , dans le volume du milieu ;
- Dans le deuxième cas d'étude, la position initiale du capteur est changée, d'un déplacement plus ou moins important.

Dans chaque configuration, la variabilité de la corrélation de coda acoustique est étudiée, afin de montrer que cette modalité est utilisable à des fins de SHM.

## 3.2.2.1 Suivi basé sur les modes

Les caractéristiques spectrales sont des paramètres particulièrement pertinents pour suivre l'état de santé d'une structure puisqu'elles découlent de sa géométrie. Un bémol sur cette approche est qu'en l'absence d'un modèle précis du milieu, il n'est pas possible d'identifier le milieu à proprement parler, seulement de suivre son évolution. Cela reste une modalité robuste, fiable et par conséquent très largement exploitée. La très grande majorité des analyses spectrales sur des bâtiments utilise des accéléromètres pour capteurs, permettant d'exploiter les sollicitations mécaniques de l'environnement sur la structure comme sources d'opportunité. Dans cette section, nous proposons de rendre compte du changement de la géométrie du milieu à partir des caractéristiques spectrales acoustiques du milieu.

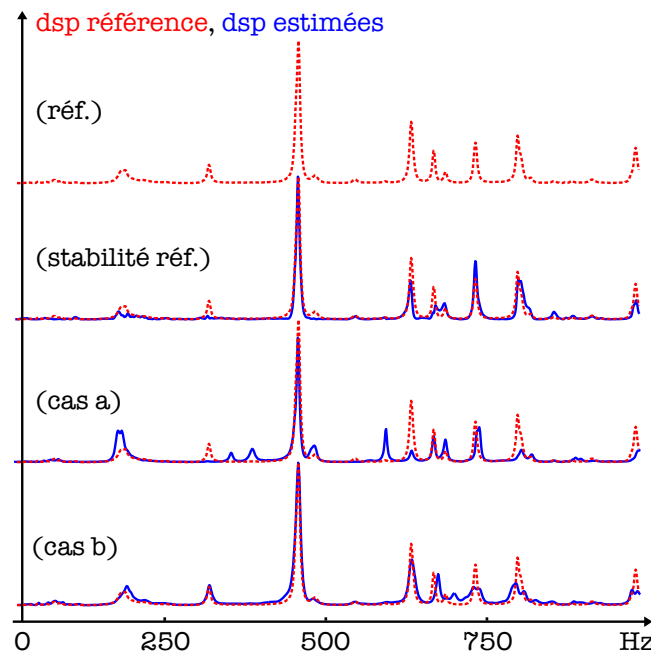


FIGURE 3.6 – Spectre de Welch d'une coda dans le cuboïde. La référence est comparée aux auto-spectres évalués dans le cas a (variation de la géométrie du milieu) et dans le cas b (déplacement du capteur de 30 cm).

La densité spectrale est estimée en calculant l'auto-spectre de Welch du champ mesuré par le capteur, sur une bande passante de 1 kHz. Le champ mesuré est la coda acoustique d'un évènement impulsionnel généré par un choc mécanique sur le milieu. Les résultats sont présentés dans la Fig. 3.6. Le spectre de référence est calculé à partir de l'ensemble des données disponibles et est affiché en pointillés. Sa représentabilité et sa robustesse sont validées par l'affichage de son estimation à partir de seulement 1/10 de la quantité de données disponibles et à partir d'un jeu de données issu d'une campagne de mesure différente. Les cas d'étude conduisent à la comparaison qualitative des auto-spectres avec la référence.

Premièrement, il apparaît que la référence est estimée de manière robuste, offrant une fonctionnelle fiable à partir de laquelle le suivi du milieu peut être effectué. Deuxièmement,

le changement de géométrie étudié dans le cas 1 correspond à une variation relative du volume du milieu de 0.16%. Malgré tout, la densité spectrale estimée est suffisamment changée pour identifier une évolution. La répartition de l'énergie dans les modes est sensiblement différente, faisant apparaître de nouveaux modes, tout en préservant certaines composantes très caractéristiques du milieu, comme le mode à 413 Hz, premier mode "plein" du milieu :

$$\omega_1 = \frac{q_{1,1,1}}{2\pi} = \frac{\sqrt{c_0^2 k_{1,1,1}^2 - d_{1,1,1}^2}}{2\pi} = 413. \quad (3.7)$$

Finalement, la position des modes n'est pas déterminée par la position du capteur et ainsi, il est normal de visualiser le même spectre dans le cas 2 où le capteur a été déplacé de quelques centimètres. En revanche, la répartition de l'énergie dans le spectre dépend de la position de la sonde, et c'est ainsi que nous expliquons les faibles variations observées.

Le résultat présenté est qualitatif et une étude quantitative complémentaire aurait pu être menée en appelant un algorithme de peak-picking pour dénombrer et localiser les modes. C'est une tâche laissée en perspective de mes travaux, et dont l'intérêt ne sera justifié, comme pour le noeud de mesure, qu'en regard d'un cahier des charges concret. Pour conclure sur le suivi de modes par la modalité acoustique, il a été montré que la construction d'une référence robuste et fiable est possible. Par ailleurs, nous avons réussi à évaluer l'impact sur cette référence de deux cas d'étude, correspondant à une variation de la géométrie du milieu et une variation de la position du capteur. Ces deux cas influencent le spectre de manière différente et il est ainsi possible de les classifier/discriminer.

### 3.2.2.2 Suivi basé sur la corrélation de Green

Dans cette dernière section, une nouvelle méthode est proposée pour le SHM et s'appuie sur le calcul de l'auto-corrélation de codas acoustiques. Cette corrélation est reliée à la corrélation de Green du milieu de propagation par une relation de filtrage, et dépend donc des paramètres physiques du milieu, comme la vitesse de propagation de l'onde, mais également de sa géométrie et de la position du point de mesure dans le milieu. Son intérêt vis-à-vis du calcul de la densité spectrale est une complexité réduite premièrement, mais aussi la possibilité de localiser les changements dans le milieu. En effet, les changements de la forme d'onde de la corrélation de Green reflètent un changement des paramètres sus-cités et la position de ces changements informe également sur leur localisation.

L'approche est similaire à celle de pour la section précédente, où les deux cas d'étude sont confrontés à une corrélation de Green de référence. Nous avons montré, à d'autres endroits dans ce manuscrit, que la fonction de Green découle de la corrélation de Green et que la robustesse de son estimation découle par conséquent de celle de la corrélation de Green. En revanche, la transformation induite par l'identité de Ward (trois dérivations notamment) amènent des performances d'estimation malgré tout légèrement moins bonne que pour la corrélation de Green. Ainsi, le choix est fait de ne mener l'étude que sur la corrélation de Green dans cette section. Il est cependant clair que l'accès à des paramètres tels que la distances aux murs ou

la distance inter-capteurs peut être obtenu par l'estimation de la fonction de Green, mais si c'était le coeur des travaux présentés dans le chapitre 2, ce n'est plus le cas à présent. Il s'agit au contraire de limiter la complexité algorithmique tout en fournissant un critère d'évaluation de l'état de santé d'une structure qui soit fiable et robuste.

Notons qu'à présent, un seul capteur est employé. La quantité d'intérêt que nous souhaitons reconstruire est donc l'auto-corrélation de Green  $C_g(t, \mathbf{x}, \mathbf{x})$ . L'auto-corrélation de référence est calculée à partir de l'ensemble des données disponibles, soit 10 ko, à partir desquels 25000 segments de coda ont été extraits (overlap autorisé entre les segments). Les autres auto-corrélations n'ont été estimées qu'à partir de 100 segments sans overlap, soit au maximum  $100 \times 44100/8 = 6$  ko de données. Elles sont alors comparées à la corrélation de référence dans la Fig. 3.7 pour les deux cas d'étude. Comme attendu, la corrélation estimée est robuste et est exploitable à des fins de monitoring. De manière très similaire à l'étude spectrale de la section précédente, un changement de la géométrie du milieu a un impact plus local sur la corrélation de Green que le déplacement du capteur, offrant de nouveau la capacité de discriminer la nature d'un changement. Dans la table adjointe à la Fig. 3.7, l'erreur quadratique moyenne entre la corrélation de référence et celle des cas d'étude est affichée, quantifiant la robustesse de la référence et l'aisance à détecter un changement dans la corrélation de Green.

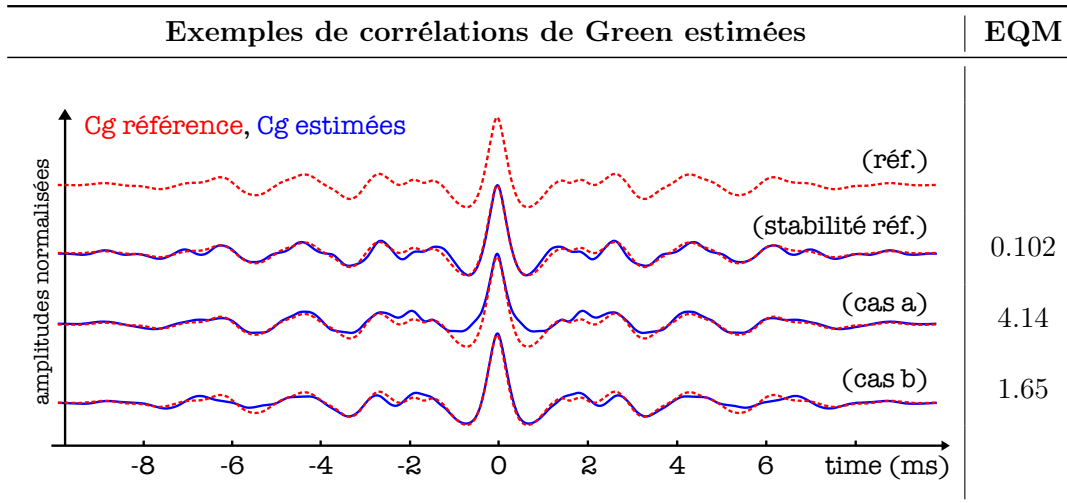


FIGURE 3.7 – Estimations de la corrélation de Green dans divers contextes.

Pour compléter les résultats expérimentaux, nous proposons une étude par la simulation de la variation de la forme d'onde de la corrélation de Green, relativement à la variation de la position du point de mesure. Nous exploiterons pour ce faire l'expression modale analytique de la corrélation de Green dans un parallélépipède rectangle. Un point de mesure est modélisé par une variable spatiale dont la valeur initiale est une position de référence. Ce point de mesure subit une perturbation additive gaussienne d'écart-type contrôlé  $\sigma$ , amenant à une nouvelle expression de l'auto-corrélation de Green. Les résultats présentés dans la Fig. 3.8 illustrent l'EQM sur la corrélation de Green en fonction de l'amplitude de cette perturbation. Les statistiques sont calculées sur 1000 itérations du procédé et sont représentées en fonction de  $\sigma/L$ , où  $L = V^{1/3} = 0.8$  m est la longueur caractéristique du milieu.

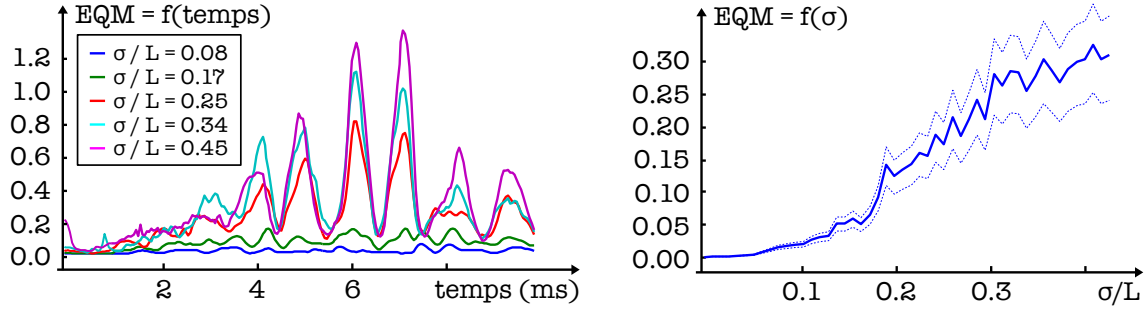


FIGURE 3.8 – Variations de la corrélation de Green avec le déplacement du point de mesure.

Ainsi, couplée à un algorithme de détection, il est sensé d'utiliser la corrélation de Green acoustique d'un milieu de propagation pour détecter et caractériser deux types de variations du système : la géométrie du milieu et la position du capteur.

Dans ce chapitre, le protocole d'identification acoustique dans un contexte passif d'estimation a été mis en oeuvre dans des scénarios crédibles et réalistes de SHM. L'algorithme d'identification passive a été adapté en combinant des techniques empruntées au monde du traitement du signal et du SHM pour réduire sa complexité algorithmique et la quantité de données manipulée. Les modifications apportées ont pour vocation de rendre embarquable le savoir-faire autour de l'identification d'un système dans un contexte passif d'estimation, à travers la modalité acoustique, ce qui représente encore à l'heure actuelle un véritable challenge technologique.

### 3.3 Ouvertures vers un identificateur passif

#### 3.3.1 Vers un noeud de mesure API

Un premier pas est à présent fait pour décrire ce que serait un noeud de mesure embarqué réalisant le suivi de l'état de santé d'une structure dans un contexte acoustique passif d'estimation. Ce projet est baptisé API (acoustic passive identification) et sera illustré dans des scénarios réalistes dans la section suivante 3.2, intégrant notamment l'estimation de fonctions de corrélation à partir de champs quantifiés sur 1 bit. Les spécificités fonctionnelles du noeud API sont liées à la modalité acoustique et plus précisément aux efforts portés pour rendre réaliste l'utilisation de cette modalité sur un noeud de mesure embarqué autonome. Quelques spécificités fonctionnelles :

- **Réveils** : L'acoustique audible requiert une fréquence d'échantillonnage élevée, qui conduit nécessairement à une quantité de données importante qui peut, si l'acquisition n'est pas contrôlée par des règles de réveil judicieuses, rapidement devenir non compatible avec la capacité de traitement d'un système embarqué. Pour répondre à cette problématique, il apparaît judicieux d'imposer le réveil ponctuel, commandé ou programmé du noeud de telle sorte que la durée de l'état d'éveil soit bien plus courte que celle de veille. Par

exemple, une évaluation journalière de l'état de santé d'une structure peut se concrétiser par le réveil quotidien du noeud pendant 10 secondes.

- **Signaux utiles à acquérir** : Les signaux utiles à l'identification du système étudié sont des codas acoustiques. Ces dernières peuvent être extraites des queues réverbérées de réponses impulsionnelles. Ainsi, durant la période d'éveil du noeud, le réveil du micro-contrôleur et l'acquisition peuvent être déclenchés par un étage analogique détecteur d'impulsion acoustique. Un tel étage est présenté en Annexe C.1. Dans l'Annexe C.2, une règle est proposée pour déterminer un critère d'arrêt et décider de la fin de la mesure de la coda. Cette règle consiste en la décision d'un seuil sur le rapport signal à bruit, sachant qu'il est lié à la probabilité que le signe du champ mesuré soit le signe de la coda et non celui du bruit de mesure. Cet étage peut attaquer une entrée analogique ou un comparateur de Schmidt du micro-contrôleur, pour échantillonner le signal sur 1 bit. La consommation à la conversion sera alors diminuée d'un facteur environ 100, sur un STM32 par exemple. Une coda acoustique dure typiquement une demi seconde, ce qui limite l'acquisition à 3 ko de données dans une configuration classique. Après cette acquisition, l'étage analogique peut être éteint et l'unité de calcul peut prendre la main pour traiter les données.
- **Informations à transmettre** : Le traitement embarqué peut estimer les quantités de Green ou des paramètres qui en découlent, sachant que les ressources de calculs et de stockage sont limitées. Les données transmises peuvent être le champ acquis et/ou les paramètres estimés à partir de ces données, ce qui représente en tout quelques ko. Il est donc attendu du transfert qu'il soit rapide et contribue en ce sens à l'autonomie du noeud.

Dans la Fig. 3.9, un exemple de schéma fonctionnel est représenté. L'étage analogique est en vert et quelques caractéristiques et dimensionnement sont proposées en rouge. Les interactions entre les différents blocs sont en bleu.

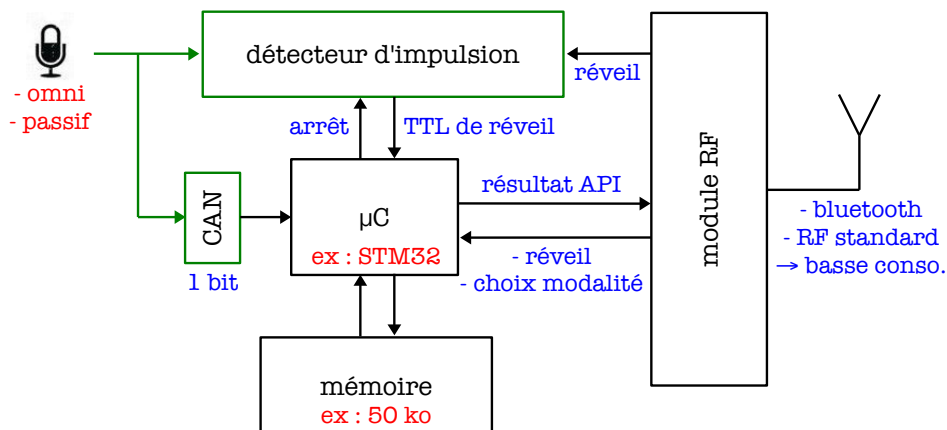


FIGURE 3.9 – Schéma fonctionnel du noeud API

Cette structure est générique et ne pourra prendre vie que face à un cahier des charges concret, lié à un besoin industriel identifié. Pour cette raison, la description fonctionnelle telle que présentée dans cette section est le niveau le plus pertinent pour décrire le noeud et sa



conception n'est pas menée plus loin dans ce manuscrit. Dans la section 3.2, des applications typiques du monde du SHM ont été discutées, mettant en oeuvre l'estimation des quantités de Green par la combinaison de la quantification 1 bit des champs et l'utilisation du décrétement aléatoire. Ces exemples sont donc réaliste par rapport à une finalité SHM et sont en adéquation avec la technologie et la capacité de traitement du noeud API.

### 3.3.2 Toujours plus loin pour l'autonomie

Je conclurai ce chapitre par une ouverture, avec une technique alternative au calcul de fonction de corrélations à partir de champs quantifiés sur 1 bits. Cette technique s'appelle le décrétement aléatoire et l'algorithme associé est sommaire et très facile à implémenter. Il jouit par ailleurs d'une très faible complexité (que des sommes), ce qui le rend particulièrement adapté au calcul embarqué. Pour ces raisons, je souhaite ouvrir la discussion sur le sujet, identifier son potentiel et les verrous à travailler.

C'est à l'époque des prémices de l'identification passive, dans les années 70, que Cole explique comment utiliser les vibrations propres d'une structure afin d'estimer des paramètres structuraux, tels que les modes propres. Voir l'article fondateur [17]. Il présente la technique du décrétement aléatoire, qui permet de déduire les résonances d'un système à partir de sa réponse à des sollicitations aléatoires. Cette technique, introduite pour caractériser des pièces de moteurs de fusées pour la NASA, ne nécessitait pas l'emploi de transformées rapides (FFT), encore peu popularisées à l'époque, et était parfaitement adaptée à la mesure en fonctionnement de l'appareil. Cette technique, appelée décrétement aléatoire, s'appuyait donc déjà sur l'écoute du bruit ambiant de la structure. Vandiver lui donne un cadre théorique (variables continues) et montre qu'il fournit une estimation de l'auto-corrélation du système tandis que J. Antoni [18] en propose une version discrétisée, adaptée pour décrire les signaux numériques. Brincker et Krenk ont généralisé la notion de décrétement aléatoire à deux variables, montrant alors que pour deux variables aléatoires gaussiennes  $X$  et  $Y$  de fonction d'inter-corrélation  $C_{XY}(t)$ , le décrétement aléatoire s'écrit :

$$D_{XY}(t) = \mathbb{E} [X(t_0 + t) | Y(t_0) = a, Y'(t_0) = b] = a \frac{C_{XY}(t)}{\sigma_Y^2} - b \frac{C'_{XY}(t)}{\sigma_{Y'}^2}. \quad (3.8)$$

où  $\bullet'$  est la dérivée temporelle de  $\bullet$ .

Autrement dit, le décrétement aléatoire est la moyenne statistique des réalisations de  $X$  de telle sorte à vérifier certaines conditions de déclenchement sur  $Y$ . Deux conditions de déclenchement interviennent et selon leur choix, le décrétement aléatoire permet d'estimer des quantités différentes. Des études complémentaires sur le décrétement aléatoire sont consultables, voir par exemple [19], montrant entre autre que l'estimateur  $\hat{D}_{XY}(\tau)$  :

$$\hat{D}_{XY}(t) = \frac{1}{N} \sum_{n=1}^N (x_n(t_0 + t) | y_n(t_0) = a, y'_n(t_0) = b) \quad (3.9)$$

est non biaisé et qu'il a une variance qui décroît linéairement avec  $N$ .

Dans l'Eq. (3.8), le terme  $aC_{XY}(t)/\sigma_Y^2$ , correspond pour Cole à la corrélation du bruit ambiant sismique du moteur de la fusée. Dans notre cas, il s'agirait d'une coda acoustique. Nous savons ainsi, après tout ce qui a été dit dans ce manuscrit, que cette corrélation estime la corrélation de Green du milieu de propagation. En disant ceci, nous venons d'expliquer pourquoi il est possible d'estimer les modes du milieu en appliquant le décréement aléatoire. Pour ne conserver que ce terme de corrélation, les conditions de déclenchement à appliquer consistent à par exemple imposer  $a \geq 0$  et à ne mettre aucune condition sur  $b$ , comptant ainsi sur le fait que statistiquement  $\mathbb{E}[b] = 0$ . En revanche, nous savons également ce que représente le deuxième terme  $-bC'_{XY}(t)/\sigma_Y^2$  ! En reprenant l'exemple de l'oscillateur amorti où la dissipation était modélisée constante, nous savons que ce terme permet d'estimer la réponse impulsionnelle de l'oscillateur ! Pour les codas sismiques, c'est également ce modèle de dissipation qui est employé, permettant donc d'estimer la fonction de Green du milieu de propagation directement par l'application du décréement aléatoire ! Pour ce faire, des conditions de déclenchement telles que  $a = 0$  et  $b \geq 0$  permettent d'y accéder.

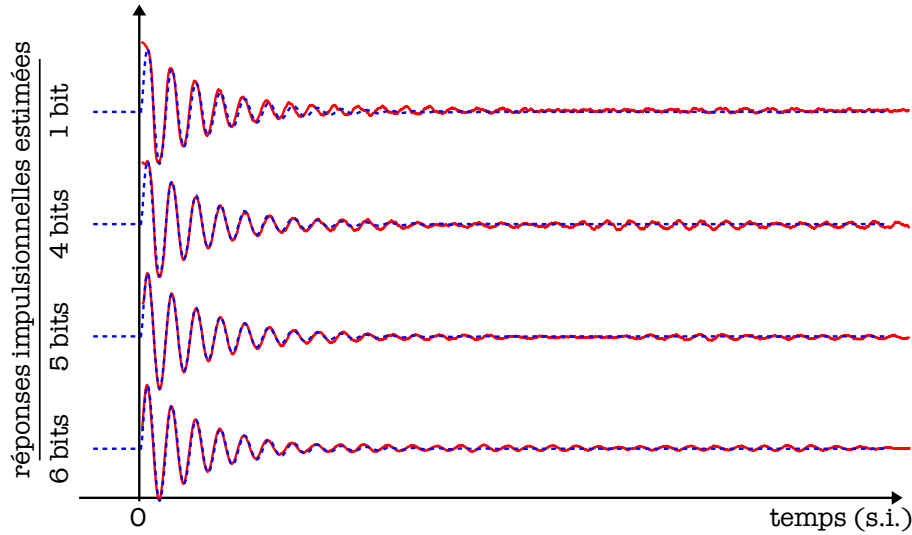


FIGURE 3.10 – Identification de l'oscillateur amorti par la combinaison du décréement aléatoire avec la quantification sur  $\{1, 4, 5, 6\}$  bits des signaux. La même quantité de données est utilisée pour chaque estimation et la fonction  $f$  est l'identité.

Dans l'Annexe E, le décréement aléatoire est comparé à l'approche d'identification passive classique de l'oscillateur amorti. Un exemple d'implémentation est proposée (en langage python) pour montrer à quel point sa complexité est faible. Notons que le décréement aléatoire est un corrélateur hybride dans le sens où le signal utilisé pour déterminer les origines des segments peut être réduit à son signe, c'est-à-dire à sa quantification sur 1 bit. Il offre donc naturellement un compromis en terme de complexité et de quantité de données manipulée. Pour ouvrir la discussion et terminer ainsi ce chapitre, je suggère l'étude de la combinaison du décréement aléatoire sur des signaux plus ou moins fortement quantifiés. Nous pouvons imaginer estimer deux quantités. L'inter-corrélation vraie normalisée des champs Gaussiens

$x(t)$  et  $y(t)$  serait obtenue à partir des champs quantifiés sur 1 bit par l'expression suivante :

$$\frac{\mathbb{E}[x(t)y(t-\tau)]}{\sqrt{\mathbb{E}[x^2(t)]\mathbb{E}[y^2(t)]}} = f\left(\frac{\pi}{2}\mathbb{E}[\text{signe}(x(t_0+\tau))|\text{signe}(y(t_0)) \geq 0]\right) \quad (3.10)$$

où  $f$  est une fonction à déterminer, tel que le sinus pour la corrélation de champs Gaussiens quantifiés sur 1 bit. Finalement, la dérivée normalisée de l'inter-corrélation des champs Gaussiens  $x(t)$  et  $y(t)$ , peut être obtenue par :

$$f\left(\frac{\pi}{2}\mathbb{E}[\text{signe}(x(t_0+\tau))|\text{signe}(y(t_0)) = 1, \text{signe}(y'(t_0)) = 1]\right) - \frac{\mathbb{E}[x(t)y(t-\tau)]}{\sqrt{\mathbb{E}[x^2(t)]\mathbb{E}[y^2(t)]}} \quad (3.11)$$

Pour conclure, je laisse ouverte l'interprétation sur les résultats obtenus pour l'identification de l'oscillateur amorti par cette méthode, Fig. 3.10, en attirant toutefois l'oeil sur le comportement au temps nul.

# Références

- [1] J. H. VAN VLECK et D. MIDDLETON, “The Spectrum of Clipped Noise”, *Proceedings of IEEE*, t. 54, n° 1, p. 2–19, 1966 (cf. p. 65).
- [2] S. S. WOLFF, J. B. THOMAS et T. R. WILLIAMS, “The Polarity-Coincidence Correlator : A Nonparametric Detection Device”, *IEEE Transactions on Information Theory*, n° 3, p. 5–9, 1962 (cf. p. 65).
- [3] H. BERNDT, “Correlation Function Estimation by a Polarity Method Using Stochastic Reference Signals”, *IEEE Transactions on Information Theory*, t. 14, n° 6, p. 796–801, 1968 (cf. p. 65).
- [4] A. K. GUPTA, “Estimation of a Quadrature Cross-Correlation Detector by Analog, Polarity-Coincidence and Relay Methods”, *IEEE Transactions on Information Theory*, t. 30, n° 6, p. 859–865, 1984 (cf. p. 65).
- [5] M. C. CHENG, “The Clipping Loss in Correlation Detectors for Arbitrary Input Signal-to-Noise Ratios”, *IEEE Transactions on Information Theory*, t. 14, n° 3, p. 382–389, 1968 (cf. p. 65).
- [6] S. GEIRHOFER, L. TONG et B. M. SADLER, “Moment Estimation and Dithered Quantization”, *IEEE signal processing letters*, t. 13, n° 12, p. 752–755, 2006 (cf. p. 65).
- [7] E. RINGOT, “RDM - calcul des ouvrages”, Université Paul Sabatier Toulouse 3, rap. tech., 2012 (cf. p. 67).
- [8] W. S. TORGERSON, “Multidimensional scaling : theory and method”, *Psychometrika*, t. 17, 1952 (cf. p. 70).
- [9] DRINEAS, “Distance Matrix Reconstruction from incomplete distance information for sensor network localization”, ENALAB, rap. tech., 2006 (cf. p. 70).
- [10] J. B. KRUSKAL, “Multidimensional Scaling by optimizing goodness of fit to a nonmetric hypothesis”, *Psychometrika*, t. 29, 1964 (cf. p. 70).
- [11] C. DAVIS et W. M. KAHAN, “Some new bounds on perturbation of subspaces”, *Bulletin of American Mathematical Society*, t. 75, 1969 (cf. p. 71).
- [12] M. CARMONA, “Identification passive des milieux. Application à la reconstruction de géométries dans les réseaux de capteurs et au diagnostic des structures”, thèse de doct., Université de Grenoble, 2008 (cf. p. 71).
- [13] M. CARMONA et O. J. J. MICHEL, *Procédé de détermination de paramètres de disposition absolue d’obets, programme d’ordinateur associé et installation de contrôle*, 2010 (cf. p. 71).
- [14] T. F. COX et M. A. A. COX, *Multidimensional scaling*, CHAPMAN et HALL, éd. 2001 (cf. p. 71).
- [15] DING, *Sensor Network Localization, Euclidean Distance Matrix, Completion and Graph Realization*. 2008 (cf. p. 71).
- [16] C. de BOOR, *A practical guide to splines*, Springer. 1978 (cf. p. 72).

- [17] H. A. COLE, “On-line failure detection and damping measurement of aerospace structures by random decrement signatures”, rap. tech., 1973 (cf. p. 78).
- [18] J. ANTONI, “Blind identification of ARMA systems using the discrete random decrement”, dans *European Signal Processing Conference*, Toulouse, 2002 (cf. p. 78).
- [19] P. KOPFF, “Renaissance d’une technique de l’analyse modale d’avant les transformées rapides : le décrétement aléatoire”, *Recherche - Traitement du signal*, t. 8, p. 345–354, 1974 (cf. p. 78).

# Estimation bayésienne de la corrélation de Green

Dans les travaux exposés précédemment, la corrélation de Green, estimée par la corrélation de coda acoustique, s'est illustrée dans un contexte applicatif concret. En réaction aux conclusions des sections précédentes, nous proposons à présent de construire un nouveau cadre d'estimation de paramètres, qui sera Bayésien, dans lequel nous proposerons un modèle pour la corrélation de Green. L'identité de Ward sera alors « court-circuitée » du protocole d'estimation de paramètres dans un contexte passif, évitant ainsi la propagation des incertitudes. Une approche construite sur un algorithme de type Chaîne de Markov - processus de Monte Carlo sera suivie pour estimer des densités de paramètres d'intérêt, tels que les temps de propagation  $t_i$ . Ces densités seront estimées *a posteriori*, c'est-à-dire en s'aidant des données à disposition et en capitalisant sur la modélisation stochastique des paramètres du modèle d'observation.

## Sommaire

<b>4.1</b>	<b>Modèle pour la corrélation de Green . . . . .</b>	<b>85</b>
<b>4.2</b>	<b>Modélisation bayésienne et algorithme d'estimation . . . . .</b>	<b>86</b>
4.2.1	Modèle d'observation et objectif . . . . .	86
4.2.2	Priors sur les paramètres du modèle . . . . .	88
4.2.3	Loi cible et sa marginalisation . . . . .	90
4.2.4	Algorithme MCMC . . . . .	91
4.2.5	Estimation de paramètres . . . . .	95
4.2.6	Illustrations . . . . .	96
<b>4.3</b>	<b>Raffinements algorithmiques et ouvertures . . . . .</b>	<b>101</b>
4.3.1	Evolution du nombre de morceaux . . . . .	101
4.3.2	Filtrage des données . . . . .	106
4.3.3	Conclusion . . . . .	107



## 4.1 Modèle pour la corrélation de Green

Dans un premier temps, nous introduisons le modèle polynomial par morceaux qui permet de décrire la corrélation de Green d'un milieu visco-acoustique. Le modèle d'observation et ses spécificités seront ensuite définis, pour présenter le cadre d'estimation bayésienne.

Pour proposer un modèle de la corrélation de Green, nous combinons deux résultats :

- L'identité de Ward visco-acoustique ;
- Le modèle de propagation hautes fréquences de la fonction de Green.

Tout d'abord, rappelons que l'identité de Ward relie par un lien fonctionnel la corrélation de Green  $C_g(t, \mathbf{x}, \mathbf{y})$  et la fonction de Green  $g(t, \mathbf{x}, \mathbf{y})$  d'un milieu visco-acoustique :

$$\alpha^2 \frac{\partial^3}{\partial t^3} C_g(t, \mathbf{x}, \mathbf{y}) \approx \text{Odd}(g(t, \mathbf{x}, \mathbf{y})). \quad (4.1)$$

Deuxièmement, en régime de propagation hautes fréquences, nous avons précisé qu'un modèle réaliste pour la fonction de Green est :

$$g(t, \mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\infty} \alpha_i \delta(t - t_i). \quad (4.2)$$

Ainsi, afin d'obtenir un modèle pour la corrélation de Green, la stratégie proposée ici est d'intégrer 3 fois l'identité de Ward. Ce faisant, nous montrons dans l'Annexe D.1 que la corrélation de Green visco-acoustique est polynomiale d'ordre 2 par morceaux :

$$C_g(t_{i+1} \geq t \geq t_i, \mathbf{x}_1, \mathbf{x}_2) = a_i t^2 + b_i t + c_i. \quad (4.3)$$

Dans la Fig. 4.1, nous illustrons ce modèle. La fonction de Green visco-acoustique est en rouge, estimée par la courbe bleue qui est issue de l'application de l'identité de Ward sur la corrélation de coda acoustique, en noir. La stratégie proposée consiste à inverser l'identité de Ward. Par intégration, les temps de propagation  $t_i$  modélisés par des diracs deviennent des instants de ruptures entre des morceaux de polynômes.

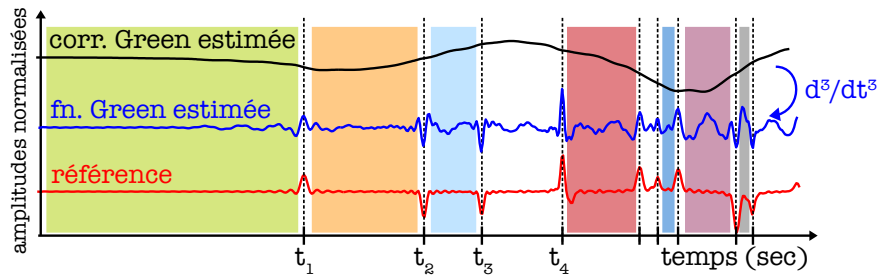


FIGURE 4.1 – Identité de Ward

Dans un contexte passif d'estimation, la coda est corrélée sur un horizon  $T$ . Sur ce support temporel, l'objectif de l'algorithme que nous souhaitons développer devra identifier le nombre



de morceaux de polynômes et localiser les ruptures (information géométrique sur le milieu de propagation), ainsi qu'estimer les coefficients des polynômes (information sur l'absorption des interfaces). Pour ce faire, nous suivrons et adapterons l'approche d'estimation bayésienne proposée par Sanquer [1, 2]. Ces travaux proposent en effet une solution particulièrement adaptée au problème rencontré.

## 4.2 Modélisation bayésienne et algorithme d'estimation

### 4.2.1 Modèle d'observation et objectif

La corrélation de Green est estimée par la corrélation de coda acoustiques. En régime de propagation hautes fréquences dans un milieu visco-acoustique, la corrélation de Green est modélisée par un polynôme d'ordre 2 par morceaux, auquel s'ajoute une nuisance d'estimation, de telle sorte qu'une expression matricielle de la fonctionnelle échantillonnée quantifiée est :

$$\mathbf{y} = \mathbf{S}\boldsymbol{\beta} + \mathbf{w} \quad (4.4)$$

où

$\mathbf{y}$	$\in \mathbb{R}^N$	observations, $C_g$ estimée sur $N$ échantillons
$\mathbf{S}$	$\in \mathbb{R}^{N \times 3(K+1)}$	matrice de design avec $K$ ruptures valant $\mathbf{t} = t_1, t_2, \dots, t_K$
$\boldsymbol{\beta}$	$\in \mathbb{R}^{3(K+1)}$	coefficients polynomiaux des $K + 1$ morceaux
$\mathbf{w}$	$\in \mathbb{R}^N$	résidu d'estimation,

Il s'agit donc d'un modèle linéaire générique, que nous allons maintenant spécifier pour notre problème.

#### 4.2.1.1 Matrice de design

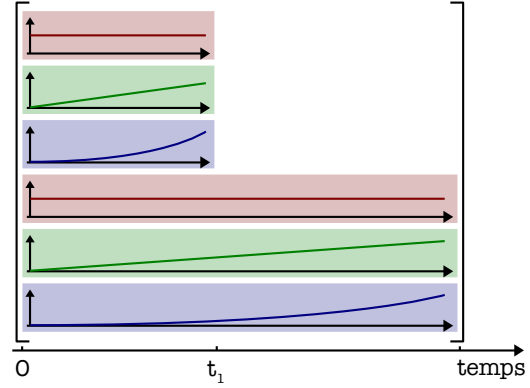
Comme exposé dans l'Annexe D.1, la corrélation de Green est construite par la superposition des polynômes d'ordre 2 par morceaux correspondant à l'intégration de chacun des temps de propagation. La matrice de design contient l'armature de cette modélisation polynomiale par morceaux : elle présente une base pour les coefficients qui décrivent le modèle. Ainsi, pour chaque morceau, la matrice est composée d'un vecteur constant, linéaire et quadratique. Une fois associée à la matrice des coefficients  $\boldsymbol{\beta}$ , chaque vecteur de la base se voit attribué un poids et leur sommation forme la corrélation complète. Un exemple de matrice de design est présenté dans la Fig. 4.2, avec une seule rupture.

La matrice de design traditionnellement utilisée pour représenter une spline par morceau est diagonale par bloc, indiquant que les morceaux ne se recouvrent pas. Elle pourrait être employée ici, mais conduirait à une perte d'information. En effet, la matrice de design proposée tient compte de ce que physiquement, les morceaux se superposent. Préserver cette forme permet ainsi d'estimer les coefficients de chacun des morceaux. C'est une spécification sans

**Légende** (Figure 4.2)

Matrice de design dans le cas où deux morceaux constituent la corrélation de Green, estimée sur une longueur  $T$ . La matrice contient un vecteur constant, linéaire et quadratique pour ces deux morceaux, permettant sa modélisation polynomiale d'ordre 2 par morceaux. La rupture entre les deux morceaux est localisée au temps de propagation  $t_1$ , c'est-à-dire le temps de propagation inter-capteur. Cette matrice conviendrait pour deux configurations :

- (1) milieu de propagation infini ;
- (2) milieu de propagation fini et  $T \leq t_2$

FIGURE 4.2 – Matrice de design  $\mathbf{S}^t$ .

conséquence sur la convergence de l'algorithme (le conditionnement de la matrice de design obéit aux mêmes contraintes que la matrice de design traditionnelle). En revanche, l'intérêt pour l'estimation est importante, car c'est le seul moyen de former un estimateur pour les amplitudes  $a_i$  des temps de propagation, porteurs d'informations sur la géométrie du milieu de propagation.

L'objectif de la stratégie d'estimation qui sera exposée dans la suite est de parvenir à estimer des paramètres de ce modèle, comme les temps de propagation  $\mathbf{t}$  ou les coefficients polynomiaux  $\beta$ .

**4.2.1.2 Bruit de mesure et vraisemblance du modèle**

Le modèle d'observation présenté ici s'intègre dans une classe de problème plus générale, celle de l'estimation de fonctions de corrélations. Ces dernières présentent des spécificités qui imposent de revisiter certaines hypothèses classiques de modélisation. Les estimateurs de corrélations présentent des biais et des variances dont le comportement asymptotique (lorsque la taille des signaux tend vers l'infini) est généralement connu ; le bruit d'estimation doit en tenir compte. De plus, la séquence additive de bruit d'estimation n'est pas nécessairement homoscédastique : en effet pour une estimation de corrélation, ses statistiques sont, d'une manière générale, dépendante de leur position dans le temps. Ici, le bruit d'estimation  $\mathbf{w}$  est donc modélisé par une séquence blanche de  $N$  échantillons normaux centrés, dont la variance est une fonction affine du temps. C'est une approximation au premier ordre qui conduit à :

$$\mathbf{w}[n] \sim \mathcal{N}\left(0, \sigma^2 \frac{n}{f_s}\right), \quad n = 1 \dots N \quad (4.5)$$

où  $f_s$  est la fréquence d'échantillonnage.

La vraisemblance du modèle vaut :

$$\mathcal{L}(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}} \frac{1}{\det(\Sigma)} \exp\left(-\frac{1}{2} \mathbf{w} \Sigma^{-1} \mathbf{w}^t\right), \quad (4.6)$$

où  $\Sigma$  est la matrice de covariance des échantillons. Son déterminant, valant  $\prod_{n=1}^N \left( \sigma^2 \frac{n}{f_s} \right)$ .

La stratégie que nous adoptons consiste à modéliser par un comportement stochastique tous les *a priori*, faibles ou forts, que nous avons sur le modèle et les paramètres. Nous pourrions alors construire une loi *a posteriori* des paramètres d'intérêt en tenant compte des données, et de leur construire des estimateurs. Pour ce faire, la première étape consiste à identifier la vraisemblance du modèle.

Connaître la famille de loi de la vraisemblance permet également d'identifier la classe des distributions qui seront conjuguées à cette vraisemblance. La conjugaison de deux distributions permet d'identifier des lois connues à partir du produit des lois conjuguées. Dans le cadre de l'estimation bayésienne mettant en jeu un estimateur MAP, comme c'est le cas ici, l'étape intermédiaire fondamentale est d'obtenir cette loi *a posteriori* des paramètres. Pour le calcul de cette loi, utiliser des lois *a priori* conjuguées de la vraisemblance, permet d'obtenir une loi *a posteriori* de la même famille et en limitant la technicité des calculs. En effet, durant l'étape de calcul de la loi *a posteriori*, la conjugaison permettra de marginaliser une partie des paramètres du modèle. Un exemple pour illustrer la notion de marginalisation est présenté dans l'Annexe D.2.

## 4.2.2 Priors sur les paramètres du modèle

Pour terminer la modélisation, faisons état des *a priori* que nous possédons. Nous proposons de construire une hiérarchie d'inférence à deux étages. Les *a priori* sont des lois choisies pour décrire le comportement des paramètres du modèle d'observation. Le choix des lois est réalisé en fonction de la connaissance de l'utilisateur sur le comportement des paramètres et la confiance qu'il accorde à cette connaissance, mais aussi de la famille de la vraisemblance du modèle, afin de choisir des *priors* qui lui sont conjugués.

Définissons les lois *a priori* de  $\beta, g$  et  $\sigma$ .

### 4.2.2.1 $\beta$ et *g*-prior

Les coefficients polynomiaux  $\beta$  sont modélisés par une loi normale centrée. Le centrage est dû au fait que le signe des amplitudes associés aux temps de propagation (les premiers) est vraisemblablement équiprobable ; par intégration, les coefficients de chaque polynôme par morceaux qui sont déduits de ces amplitudes ne favorisent donc aucun signe. La décroissance de l'amplitude du champ acoustique se propageant, intégrant atténuation géométrique et dissipative, peut se mettre sous la forme d'une exponentielle décroissante, ce qui n'est pas en contradiction avec le choix d'une loi normale, par ailleurs conjuguée de la vraisemblance.

Un *g*-prior conjugué à la vraisemblance est employé pour paramétrer la variance de  $\beta$  :

$$\beta | \mathbf{t}, \sigma^2, g \sim \mathcal{N} \left( 0, g [\mathbf{S}^t \Sigma \mathbf{S}]^{-1} \right) \quad (4.7)$$

Ce choix est motivé par le fait que le g-prior conduit, comme le montre Sanquer [1, 2], à une loi postérieure qui est invariante par transformation linéaire de la base de régression (les polynômes d'ordre 0, 1, 2). Ainsi, l'algorithme reste applicable à des données de dynamiques différentes, et ne dépend en réalité pas de la base de polynômes choisie. Dans ses travaux, Sanquer interprète et démontre cette propriété. En particulier, nous citons ici son interprétation de  $g$ . Considérons une matrice de design  $\mathbf{S}$  donnée et évaluons :

$$\mathbb{E}[(\mathbf{S}\boldsymbol{\beta})^t \mathbf{S}\boldsymbol{\beta}] = \mathbb{E}[\text{tr}((\mathbf{S}\boldsymbol{\beta})^t \mathbf{S}\boldsymbol{\beta})] = \text{tr}(\mathbf{S}^t \mathbb{E}[\boldsymbol{\beta}\boldsymbol{\beta}^t]) = \text{tr}(\mathbf{S}^t \mathbf{S} g (\mathbf{S}^t \Sigma \mathbf{S})^{-1}) \quad (4.8)$$

et donc :

$$g = \frac{\mathbb{E}[(\mathbf{S}\boldsymbol{\beta})^t \mathbf{S}\boldsymbol{\beta}]}{\text{tr}(\Sigma^{-1})} \quad (4.9)$$

Le terme  $(\mathbf{S}\boldsymbol{\beta})^t \mathbf{S}\boldsymbol{\beta}$  représente l'énergie totale du modèle tandis que la matrice de covariance  $\Sigma$  contient l'information sur l'énergie du bruit d'estimation. Ainsi, le paramètre  $g$  est interprété comme une mesure du rapport signal-à-bruit, conditionnellement à la matrice de design. Puisque  $\sigma$  rend compte du bruit de modèle, nous pouvons ainsi considérer que  $g|\sigma$  rend compte de l'énergie du signal utile, au travers du RSB.

#### 4.2.2.2 Full bayésien et priors peu informatifs

Le cadre d'estimation bayésienne hiérarchisé est construit par strates, où les paramètres du modèles sont décrits conditionnellement par d'autres. C'est donc une structure d'inférence d'information inter-paramètres, où la modélisation des *a priori* de chaque paramètre s'appuie la confiance toute relative que l'utilisateur porte à sa connaissance du modèle. Les références [3, 4, 5] fournissent une introduction complète sur le sujet.

Ce paradigme d'inférence hiérarchisée peut être mis en oeuvre de plusieurs manières et nous distinguons ici deux écoles principales, que nous avons rebaptisées :

**L'école du bonzaï :** L'approche bayésienne empirique contrôle la profondeur de l'arbre hiérarchique en choisissant des valeurs pour les paramètres dès que c'est possible. Elle s'appuie sur les données pour former des estimateurs des paramètres de la densité marginale postérieure en se basant sur les observations, par exemple par maximum de vraisemblance, par les moindres carrés, par la méthode des moments... l'intérêt de cette approche réside dans le fait qu'un paramètre fixé n'est plus à échantillonner, ce qui conduit à une réduction de la complexité hiérarchique et algorithmique. En revanche, les données sont utilisées à plusieurs reprises au cours de la chaîne d'estimation (pour choisir les paramètres puis pour estimer la densité postérieure), ce qui peut conduire à un modèle trop spécialisé sur le jeu de données (à la manière du sur-apprentissage) ;

**L'école du saule pleureur :** L'approche tout-bayésien favorise une hiérarchie à tiroirs, intégrant sans différenciation tous les niveaux de confiance, en faisant appel pour les *a priori* les plus vagues à des lois peu voire non informatives. L'algorithme repose entièrement sur des stratégies de décisions probabilistiques pour parcourir l'espace d'états,

converger et estimer les paramètres. L'intérêt de cette approche est qu'elle est générique, au détriment d'une complexité algorithmique qui peut rapidement exploser.

Dans la littérature sur la régression polynomiale par morceaux, voir par exemple [6, 7, 8, 1, 2, 9], c'est la stratégie tout-bayésien qui est retenue à l'unanimité. Les modèles proposés sont d'ailleurs généralement plus complexes que le nôtre, estimant l'ordre des polynômes, paramétrant les transitions entre les morceaux. . . En revanche, il serait inefficace de volontairement ignorer la connaissance disponible à propos de certains paramètres et l'algorithme proposé tiendra compte des spécificités dictées par la physique. Une étude complète de comparaisons de choix à propos des hyper-paramètres est proposée par Mackay dans [5]. Lorsque la confiance portée à la modélisation d'un paramètre sera basse voir nulle, son poids dans la hiérarchie de l'inférence devra être moindre. Jeffreys [10] propose une méthode pour construire de telles lois, qui réduisent voire rendent nulle la quantité d'information apportée par le paramètre, dans le modèle. Il exploite l'information de Fisher et construit des lois avec, potentiellement impropres. Le fait que les lois soient impropres (l'intégrale de la densité ne vaut pas 1 par exemple), n'est pas nécessairement un problème. Ici, ces lois impropres permettent de modéliser des paramètres avec un *a priori* faible et destinées à être marginalisées. Pour illustrer la notion de *a priori* peu informatif de Jeffrey, considérons la densité

$$p(x|\lambda) = \lambda \exp(-\lambda x) \quad (4.10)$$

qui amène à l'information de Fisher :

$$I(\lambda) = \mathbb{E} \left[ -\frac{\partial^2}{\partial \lambda^2} \log p(x|\lambda) \right] = \frac{1}{\lambda^2}. \quad (4.11)$$

La densité que Jeffreys propose est  $\pi(\lambda) \propto \sqrt{I(\lambda)} \propto 1/\lambda$ . L'information de Fisher, qui se compare à l'inverse d'une variance, est donc utilisée ici pour construire une loi impropre (son intégrale ne vaut pas l'unité).

Ici, les hyper-paramètres du modèle que sont  $\sigma^2$  et  $g$  influencent la variance de lois normales ; une loi inverse-gamma, qui est par ailleurs une loi conjuguée à la vraisemblance, est adéquate pour les modéliser. Cependant, ces paramètres forment la strate d'inférence la plus profonde de la hiérarchie et ils peuvent être modélisés par des priors peu, voir non informatifs. C'est une liberté dans la modélisation qui sera profitable, notamment pour mener des calculs de marginalisation, ce que nous abordons dans la section suivante.

### 4.2.3 Loi cible et sa marginalisation

Les paramètres du modèle sont  $\theta = [\mathbf{t}, \beta, \sigma^2, g]$ . La densité *a posteriori* du modèle est  $p(\theta|\mathbf{y})$ . Cette densité peut être marginalisée vis-à-vis de certains des paramètres, car la propagation de leur modèle dans la densité peut être explicitée et même complètement intégrée. Ces paramètres marginalisables sont appelés paramètres de nuisance, complexifiant la forme de la densité postérieure, voir [11]. Suivant cette philosophie, nous proposons de réduire la densité *a posteriori* du modèle à  $p(\mathbf{t}, g|\mathbf{y})$ . Cet état pour la densité marginalisée est le plus compact

que nous puissions obtenir de manière exacte, ne faisant intervenir que les deux seuls paramètres essentiels au problème : la position des ruptures entre les morceaux et une évaluation du rapport signal à bruit.

L'astuce de calcul mise en oeuvre dans l'étape de marginalisation consiste à exploiter la loi de Bayes pour déplier la hiérarchie d'inférence et ainsi faire apparaître des densités de probabilités connues. Leur intégration sur l'espace complet valant l'unité, leur contribution pourra être rendue complètement transparente dans la densité postérieure, qui prendra alors une expression moins complexe.

Compte-tenu des relations d'indépendances entre les paramètres et en utilisant la loi de Bayes, nous obtenons :

$$p(\mathbf{t}, g, \boldsymbol{\beta}, \sigma^2 | \mathbf{y}) \propto \mathcal{L}(\mathbf{y} | \mathbf{t}, g, \boldsymbol{\beta}, \sigma^2) p(\mathbf{t}) p(\sigma^2) p(g) p(\boldsymbol{\beta} | \sigma^2, g). \quad (4.12)$$

La marginalisation des paramètres s'exprime par :

$$p(\mathbf{t}, g | \mathbf{y}) \propto \iiint_{\mathbb{R}^{3K+1} \times \mathbb{R} \times \mathbb{R}} \mathcal{L}(\mathbf{y} | \mathbf{t}, g, \boldsymbol{\beta}, \sigma^2) p(\mathbf{t}) p(\sigma^2) p(g) p(\boldsymbol{\beta} | \sigma^2, g) d\boldsymbol{\beta} d\sigma. \quad (4.13)$$

et le calcul est mené en Annexe D.3. La distribution *a posteriori* marginalisée obtenue est

$$p(\mathbf{t}, g | \mathbf{y}) \propto \frac{p(\mathbf{t}) p(g) (1 + g)^{-\frac{3(K+1)}{2}}}{\left[ \mathbf{y}^T \Sigma^{-1} \{ \mathbf{1}_N - M \} \mathbf{y} \right]^{\frac{N}{2}}}, \text{ où } M = \frac{g}{1 + g} \mathbf{S} [\mathbf{S}^T \Sigma^{-1} \mathbf{S}]^{-1} \mathbf{S}^T \Sigma^{-1}. \quad (4.14)$$

La densité postérieure obtenue n'a pas une forme standard de densité et ne peut être davantage marginalisée. Avec une stratégie tout-bayésien, la densité postérieure cible prend souvent une forme complexe et non standard, ne permettant que très rarement l'emploi des estimateurs bayésiens classiques de types maximum de vraisemblance, espérance conditionnelle, etc. qui en fournissent une expression analytique. C'est manifestement le cas ici. Alors, comme bien souvent, les paramètres recherchés ( $\mathbf{t}$  ici en particulier) seront estimés à partir d'une version échantillonnée de la densité cible, obtenue par une technique à définir et présentée dans la section suivante.

#### 4.2.4 Algorithme MCMC

Afin d'échantillonner une densité de distribution, nous employons une méthode de Monte Carlo. Cette dernière est mise en oeuvre grâce à un algorithme itératif qui parcourt une chaîne de Markov, dont la densité stationnaire est la densité à estimer. Une fois la densité cible estimée, ici la densité postérieure marginalisée  $p(\mathbf{t}, g | \mathbf{y})$ , les paramètres souhaités seront estimés en allant extraire le maximum *a posteriori* (estimateur MAP). Un algorithme de type Metropolis-Hastings [12] est décrit pour résoudre le problème posé.

Une chaîne de Markov est construite avec pour densité stationnaire  $\pi(\boldsymbol{\theta})$  ; à chaque itération  $n$ , un mouvement de la chaîne propose un échantillon de la densité cible,  $\boldsymbol{\theta}_{n+1}^*$ , à partir des

paramètres courants  $\theta_n$ , selon une loi  $q(\theta_n \rightarrow \theta_{n+1}^*)$ . Celui-ci est proposé mais n'est pas nécessairement retenu ; le mouvement opéré est évalué par un score et une décision est prise accepter ou non le mouvement. Nous définissons le taux d'acceptation d'un mouvement tel que le suggère Peskun [13], par la quantité :

$$\alpha(\theta_{n+1}) = \min\left\{1, \frac{\pi(\theta_{n+1}^*)}{\pi(\theta_n)} \times \frac{q(\theta_n \leftarrow \theta_{n+1}^*)}{q(\theta_n \rightarrow \theta_{n+1}^*)}\right\}, \quad (4.15)$$

car elle génère l'estimateur pour la densité cible qui a la variance asymptotique minimale. Le mouvement est accepté avec la probabilité  $\alpha$ . Les paramètres marqués d'un \* sont les paramètres mis à jours et proposés durant le mouvement.

#### 4.2.4.1 Mouvement de mise à jour des ruptures

La densité cible ici est  $\pi(\theta) = p(\mathbf{t}, g | \mathbf{y})$ . Pour le moment, la chaîne de Markov n'est constituée que d'un seul mouvement : la mise à jour des ruptures  $\mathbf{t}$  entre les morceaux et  $g$ . Ici, parmi toutes les stratégies de mise à jour possible, il semble judicieux de commencer par mettre à jour  $\mathbf{t}$  puis  $g$ , car  $g$  est une mesure du RSB calculée conditionnellement aux positions des ruptures. Par ailleurs, nous choisissons de restreindre la mise à jour à un seul des temps par itération. En effet, de manière générale, mettre à jour plusieurs paramètres en même temps conduit à une variabilité de la rapidité de convergence de l'algorithme, puisque le nombre de configurations en sorties est d'autant plus grand qu'il y a de paramètres mis à jour.

**Légende** (Figure 4.3)

La rupture  $t_i$  à mettre à jour est choisie uniformément parmi le jeu de ruptures. La nouvelle valeur est le résultat d'un déplacement aléatoire à partir de la position précédente. Par exemple,

$$t_{i,n+1}^* | t_{i,n} \sim \mathcal{N}\left(t_{i,n}, \frac{N}{L}\right) \quad (4.16)$$

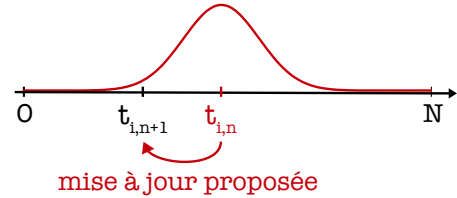


FIGURE 4.3 – Mise à jour d'une rupture.

où  $L$  sera choisi empiriquement afin de contrôler le temps de calcul et la vitesse de convergence.

Après cette mise à jour, une nouvelle matrice de design est construite. Les autres paramètres du modèle peuvent maintenant être mis à jour en conséquence, l'objectif étant d'arriver à une nouvelle estimation de  $g$ . Pour ce faire, nous emploierons un échantillonneur de Gibbs, qui est un algorithme également de type MCMC qui permet d'obtenir une séquence d'échantillons estimant une densité de distribution cible ; ici, la densité jointe de  $(g, \sigma^2, \beta)$ . Pour obtenir un échantillon de cette densité, chaque paramètre est successivement échantillonné, conditionnellement aux autres, pris à leur valeur la plus récemment obtenue ; les échantillons obtenus approximent alors la densité jointe cible. A partir d'un échantillonneur de Gibbs, des choix algorithmiques peuvent parfois être réalisés pour optimiser la convergence de l'algorithme tout en préservant la cible (la densité stationnaire de la chaîne de Markov). En guise d'exemple, citons ici le « partially collapsed Gibbs sampler », qui reposent sur 3 types de manipulations

de la succession des estimations des paramètres : la marginalisation, la permutation et la suppression des étapes qui conduisent à l'estimation de quantités intermédiaires non utilisées par la suite pour conditionner d'autres variables.

Dans le problème présenté, la densité postérieure jointe des paramètres est  $p(\mathbf{t}, \boldsymbol{\beta}, \sigma^2, g | \mathbf{y})$ . Son estimation peut donc être réalisée par un échantillonnage de Gibbs, c'est-à-dire par le tirage successif de chacun des paramètres du problème, conditionnellement aux autres. Par ailleurs, nous appliquons l'approche du « partially collapsed Gibbs sampling », voir par exemple [14], en recherchant la marginalisation de certains paramètres. L'Annexe D.4 résume le raisonnement et les calculs menés dans notre cas. L'échantillonnage choisi s'articule en commençant par  $\sigma^2$  :

$$\sigma^2 | \mathbf{y}, \mathbf{t}, g \sim \mathcal{IG}\left(\frac{n}{2}, \frac{1}{2}(\mathbf{y}^T \Sigma^{-1} \{\mathbb{1}_n - M\} \mathbf{y})\right) \quad (4.17)$$

où  $\boldsymbol{\beta}$  a été marginalisé.

Dans un deuxième temps, c'est maintenant  $\boldsymbol{\beta}$  qui est mis à jour :

$$\boldsymbol{\beta} | \mathbf{y}, \sigma^2, \mathbf{t}, g \sim \mathcal{N}\left(M \mathbf{y}, \sigma^2 \frac{g}{1+g} S[S^T \Sigma^{-1} S]^{-1}\right) \quad (4.18)$$

Et finalement, le paramètre  $g$  :

$$g | \sigma^2, \mathbf{t}, \boldsymbol{\beta} \sim \mathcal{IG}\left(A + \frac{3K}{2}, B + \frac{1}{2\sigma^2} \boldsymbol{\beta}^T S^T \Sigma^{-1} S \boldsymbol{\beta}\right) \quad (4.19)$$

où  $A = 1, B = 100$  sont fixés empiriquement, témoignant un *a priori* peu informatif sur  $g$ .

La probabilité d'acceptation du mouvement décrit est alors :

$$\alpha = \min\left\{1, \frac{p(\mathbf{t}^*, \mathbf{g}^* | \mathbf{y})}{p(\mathbf{t}, \mathbf{g} | \mathbf{y})} \times \frac{q(\mathbf{t}, g \leftarrow \mathbf{t}^*, g^*)}{q(\mathbf{t}, g \rightarrow \mathbf{t}^*, g^*)}\right\}, \quad (4.20)$$

avec

$$\frac{p(\mathbf{t}^*, \mathbf{g}^* | \mathbf{y})}{p(\mathbf{t}, \mathbf{g} | \mathbf{y})} = \left(\frac{1+g^*}{1+g}\right)^{-\frac{3(K+1)}{2}} \left[\frac{\mathbf{y}^T \Sigma^{-1} \{\mathbb{1}_N - M\} \mathbf{y}}{\mathbf{y}^T \Sigma^{-1} \{\mathbb{1}_N - M^*\} \mathbf{y}}\right]^{\frac{N}{2}} \quad (4.21)$$

et

$$\frac{q(\mathbf{t}, g \leftarrow \mathbf{t}^*, g^*)}{q(\mathbf{t}, g \rightarrow \mathbf{t}^*, g^*)} = \frac{p(t_i)}{p(t_i^*)} \times \frac{q(t_{i,n} | t_{i,n+1}^*)}{q(t_{i,n+1}^* | t_{i,n})} \times \frac{p(g | \mathbf{t})}{p(g^* | \mathbf{t}^*)} = \frac{p(g | \mathbf{t})}{p(g^* | \mathbf{t}^*)} \quad (4.22)$$

car le tirage de la rupture est uniforme parmi toutes les ruptures et la marche de mise à jour de la rupture est une loi symétrique.

L'algorithme 3 résume la stratégie adoptée pour échantillonner la densité cible. Certains paramètres comme  $N_{max}$  et  $N_{bi}$  seront précisés et expliqués en regard de quelques considérations algorithmiques.



**Algorithme 3:** Evolution de la chaîne MCMC

---

<b>while</b> $n \leq N_{max}$ <b>do</b>	<b>1</b>
- choix aléatoire de la rupture $i$ à mettre à jour	<b>2</b>
- mise à jour de la position $t_i$ de la rupture $i$	<b>3</b>
- mise à jour des autres paramètres par le « partially collapsed Gibbs sampling »	<b>4</b>
- calcul du score $\alpha$ obtenu avec la mise à jour	<b>5</b>
- mouvement accepté avec la probabilité $\alpha$	<b>6</b>
<b>if</b> <i>mouvement accepté</i> <b>then</b>	<b>7</b>
- paramètres actuels conservés	<b>8</b>
<b>if</b> $n \geq n_{bi}$ <b>then</b>	<b>9</b>
- ajouter l'échantillon à la densité cible	<b>10</b>
<b>end</b>	<b>11</b>
<b>else</b>	<b>12</b>
- paramètres précédents restaurés	<b>13</b>
<b>end</b>	<b>14</b>
<b>end</b>	<b>15</b>

---

**4.2.4.2 Considérations algorithmiques**

Après initialisation de l'algorithme MCMC, une période de chauffe de  $N_{BI}$  mouvements acceptés de la chaîne est subie avant de commencer l'échantillonnage des paramètres à proprement parler. Cette approche est une précaution initiée pour faire face aux mauvaises initialisations ; la densité estimée est, pendant une certaine période, non représentative de la distribution cible et les échantillons ne sont pas gardés en mémoire. Les mouvements acceptés pendant cette période de chauffe permettent simplement à l'algorithme de converger vers la distribution cible à estimer. Définir un critère de sortie de la période de chauffe est un problème encore ouvert. Des choix empiriques sont souvent faits, par exemple en observant la convergence des paramètres du problèmes à partir de plusieurs instances de l'algorithme, initialisées différemment. Une quantité centrale à notre problème est :  $P = \mathbf{y}^T \Sigma^{-1} \{\mathbf{1}_n - M\} \mathbf{y}$ . Dans la suite, nous observerons la convergence de cette quantité pour en déduire la durée moyenne d'une période de chauffe.

Les échantillons fournis par un échantillonneur de Gibbs forment une chaîne de Markov, dont les éléments sont corrélés de proche en proche. Une méthode de « thinning » peut être appliquée si les éléments recherchés doivent être décorrélés. Il s'agira alors de ne retenir qu'un échantillon sur  $X$  estimés. Dans l'Annexe D.5, nous proposons une méthode pour choisir  $X$  de façon à ce que le coefficient de corrélation entre les échantillons successifs d'une marche aléatoire soit constant. C'est une condition qui accroît fortement le nombre d'itérations requis et l'indépendance n'apparaît pas comme nécessaire. Nous n'appliquerons donc pas cette règle.

### 4.2.5 Estimation de paramètres

#### 4.2.5.1 Estimation des transitions et des paramètres polynomiaux

Lorsque la chaîne a réalisé un certain nombre de mouvements acceptés, il appartient à l'utilisateur de décider d'arrêter la chaîne. Il est clair que la décision s'appuie sur un compromis temps de calcul / précision de l'estimation. Dans le cadre étudié, aucune contrainte de complexité ou de temps de calcul n'est imposée, aussi les densité cibles seront estimées avec un grand nombre d'échantillons (plusieurs dizaines de milliers).

A partir de la densité cible estimée, la position des ruptures entre les morceaux de polynôme, sera estimée par l'estimateur par maximum *a posteriori* :

$$\forall i, \quad \hat{\mathbf{t}}_{i,MAP} = \arg \max_{\mathbf{t}_i} p(\mathbf{t}_i, g | \mathbf{y}) \quad (4.23)$$

au détriment de l'estimateur d'espérance conditionnelle :

$$\forall i, \quad \hat{\mathbf{t}}_{i,EC} = \mathbb{E} [p(\mathbf{t}_i, g | \mathbf{y})] . \quad (4.24)$$

Ce choix trouve une justification dans le fait que la densité cible n'est pas symétrique *a priori*, et que nous préférons estimer la position la plus vraisemblable plutôt que la position moyenne. Par ailleurs, il est permis à l'utilisateur d'estimer la densité cible par un grand nombre d'échantillon, où localiser un maximum est réalisé avec une confiance d'autant plus grande que le nombre d'échantillons augmente. Dans le cas inverse où très peu d'échantillons sont autorisés, il semble au contraire naturel de préférer estimer l'espérance mathématique de la densité plutôt que de localiser un maximum mal défini. Par ailleurs, quitte à commettre une erreur si la densité cible n'est pas mono-modale, autant que l'estimateur ait malgré tout un sens et corresponde à une probabilité max locale, ce qui est atteint avec l'estimateur MAP. Enfin, cet estimateur autorise, étant donné sa définition, à réduire voire à annuler complètement la période de chauffe de la chaîne.

#### 4.2.5.2 Reconstruction de la corrélation de Green

Dans la perspective de SHM, nous avons montré dans les sections précédentes que la forme de la corrélation de Green est une modalité valable pour identifier des changements géométriques du milieu de propagation. Or, si la position des ruptures est connue dans le modèle d'observation Eq. (4.25), reconstruire la corrélation de Green se résout par l'estimation des coefficients  $\beta$ . Par exemple, si les positions des ruptures entre les morceaux de polynômes sont estimés par maximum *a posteriori*, ainsi que les paramètres  $\sigma^2$  et  $g$ , alors le modèle d'observation devient :

$$\mathbf{y} = \mathbf{S}_{MAP} \beta + \mathbf{w}_{MAP} \quad (4.25)$$

où

$$\begin{aligned} \mathbf{y} &\in \mathbb{R}^N, \text{ sont les observations} \\ \mathbf{S}_{MAP} &\in \mathbb{R}^{N \times 3(K+1)}, \text{ est la matrice de design, formée à partir de } \mathbf{t}_{MAP} \\ \boldsymbol{\beta} &\in \mathbb{R}^{3(K+1)}, \text{ sont les coefficients polynomiaux à estimer} \\ \mathbf{w}_{MAP} &\in \mathbb{R}^N, \text{ résidu d'estimation de variance } \sigma_{MAP}^2. \end{aligned}$$

Deux stratégies s'offrent à nous pour y parvenir. De la même manière que les temps de propagation, nous pouvons isoler le MAP pour  $\boldsymbol{\beta}$ . En revanche, le lien avec le MAP pour  $\mathbf{t}$  n'est pas évident, aussi nous préférons estimer les coefficients polynomiaux en formulant un problème bien connu de régression sous contrainte, pour lequel une solution explicite existe. En effet, la modèle d'observation où la nuisance est additive et gaussienne permet de formuler un problème d'estimation convexe où la solution  $\hat{\boldsymbol{\beta}}$  vérifie :

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} \left\{ \sum_i (\mathbf{y}_i - \mathbf{S}_i^t \boldsymbol{\beta})^2 + (\Gamma \boldsymbol{\beta})^t (\Gamma \boldsymbol{\beta}) \right\}. \quad (4.26)$$

Dans l'Eq. (4.26), le critère d'attache aux données est complété par une régularisation de l'énergie des coefficients polynomiaux. Ce problème bénéficie d'une expression explicite de la solution, évitant l'emploi d'un algorithme d'optimisation :

$$\hat{\boldsymbol{\beta}} = (\mathbf{S}^t \mathbf{S} + \Gamma^t \Gamma)^{-1} \mathbf{S}^t \mathbf{y}. \quad (4.27)$$

La matrice de Tikhonov  $\Gamma$  est généralement déterminée par des méthodes ad-hoc d'apprentissage, telles que la validation croisée [15], L-curve [16]... mais est généralement difficile à interpréter. Dans le cas scalaire, le choix du paramètre de régularisation est souvent très fortement lié aux statistiques du bruit de modèle, voir par exemple [17]. Ici, nous proposons de faire de même et de choisir pour  $\Gamma$  la matrice de covariance du bruit de modèle.

#### 4.2.6 Illustrations

Dans un premier temps, l'algorithme d'estimation présenté est mis en oeuvre sur un jeu de données polynomial par morceaux avec 1 seule rupture. Cet exemple est utilisé pour caractériser le comportement de l'algorithme. Dans un second temps, l'approche est confrontée sur un jeu de données construit avec plusieurs ruptures. Une discussion à propos du choix du nombre de ruptures est alors proposée, amenant sur la section suivante du chapitre, où ce choix de modèle est injecté au sein de la chaîne de Markov.

##### 4.2.6.1 Toy problem A

Le premier jeu de données ne contient qu'une seule rupture. Le bruit ajouté au polynôme est gaussien centré de variance  $\sigma_0^2(1 + \alpha t)$ . De manière générale pour un estimateur de corrélation,

la relation  $1 \gg \alpha n T_s$  est vérifiée à moins que l'horizon d'estimation soit très grand. Ainsi,  $\alpha$  est choisi égal à 0.01. Dans la Fig. 4.4, un résultat typique est représenté, lorsque le rapport signal à bruit est plutôt favorable (0 dB).

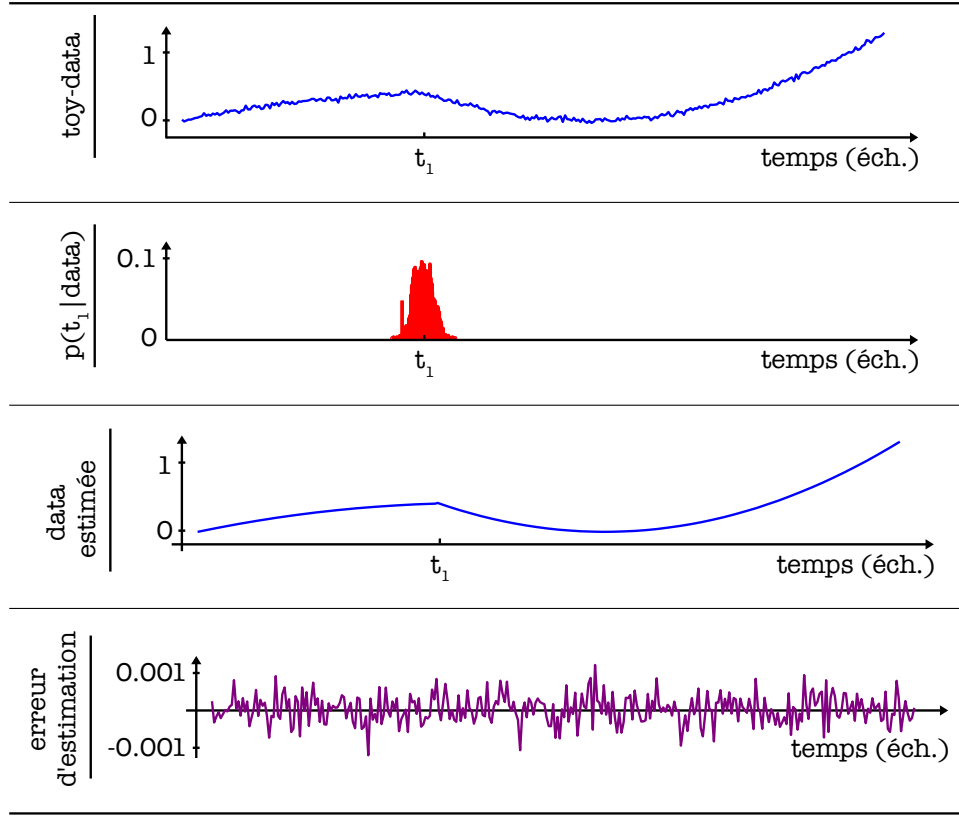


FIGURE 4.4 – Toy problem. Données brutes, densité postérieure, données régressées et erreur.

La densité  $p(t_1|data)$  est construite avec 10000 échantillons. Les données reconstruites sont estimées par application de l'estimation de *maximum a posteriori* pour estimer  $t_1$ , Eq. (4.23) puis les coefficients polynomiaux sont obtenus en résolvant le problème de régression ridge présenté 4.2.5.2. La reconstruction des données est fidèle à la réalité, l'erreur représentée en violet étant très proche du bruit additif venu perturber les données. L'erreur RMS moyenne relative entre cette erreur et le bruit de modèle est de l'ordre de 0.01%.

Dans la Fig. 4.5, quelques statistiques sont représentées afin de caractériser le comportement des densités postérieures du temps de propagation. L'erreur *root-mean-square* (RMS) est comparée à l'écart-type des densités postérieures, pour divers RSB. Les différences observées sont en partie dues à la non symétrie de la densité postérieure. Le comportement observé est que la densité postérieure s'étale très largement sur l'ensemble du support temporel de la fonctionnelle à régresser lorsque le RSB est trop bas. Ceci induit une variance d'autant plus grande que le RSB avec un maximum donné par la largeur du support. Lorsque le RSB s'améliore, la densité se concentre et permet d'obtenir des distributions du type de la Fig. 4.4. Ce comportement est exacerbé, dans un sens comme dans l'autre, lorsque le nombre d'échantillons utilisés pour estimer la densité cible augmente. En réalité, cette observation est valide en ré-

gime d'estimation non asymptotique, lorsque le nombre d'échantillon estimé reste « petit ». En pratique, sur l'exemple A,  $10^5$  échantillons permettent d'atteindre le régime asymptotique dans lequel les statistiques de la densité cible sont stabilisées. Dans ce régime et lorsque le RSB est supérieur à 5 dB, l'erreur RMS commise sur l'estimation de la position de la rupture en devient alors plus petite que la période d'échantillonnage.

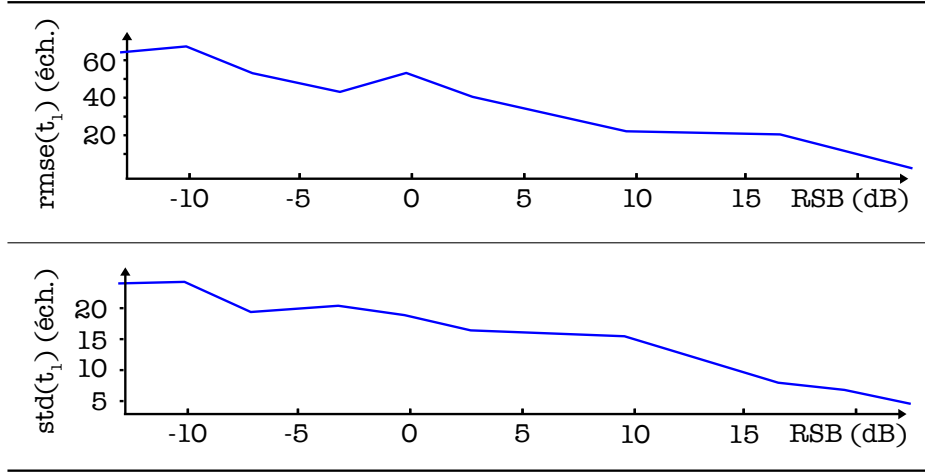


FIGURE 4.5 – Toy problem. Statistiques en fonction du RSB.

Sur l'exemple A, le nombre de mouvements requis à la chaîne pour estimer un nombre  $N$  d'échantillons donné est approximativement  $2N$ . Il ne dépend pas du RSB ni de l'horizon d'estimation. C'est un signe que la période de burn-in a été correctement mise en oeuvre. Cela montre également que la chaîne est « équilibrée », dans le sens où globalement la probabilité que le (seul) mouvement de la chaîne défini soit accepté est 0.5.

Dans la Fig. 4.6, le comportement moyen du paramètre  $P$  est représenté lorsque celui-ci est initialisé aléatoirement, ici tiré uniformément sur une plage de plusieurs décades. La trace montre l'atteinte d'un état stabilisé après un nombre  $N_{burn-in}$  d'échantillons. Ce nombre est de l'ordre de 500 sur cet exemple. En pratique, pour améliorer la probabilité que ce paramètre ait convergé, l'échantillonnage effectif de la distribution cible pourra débuter après un nombre de mouvement de la chaîne plus grand, typiquement 2 ou 3 fois. Une bonne initialisation vaut parfois mieux qu'une période de chauffe longue et dont la convergence n'est pas maîtrisée, mais dans un esprit de généralisation, il était intéressant de quantifier le comportement d'une période de chauffe générique.

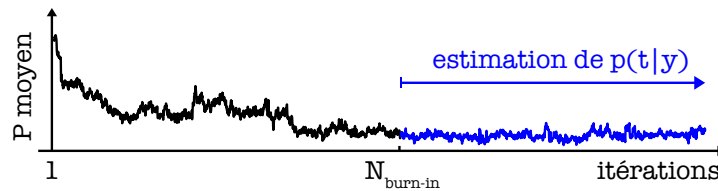


FIGURE 4.6 – Comportement moyen du paramètre  $P = \mathbf{y}^T \Sigma^{-1} \{\mathbb{1}_n - M\} \mathbf{y}$ .

## 4.2.6.2 Toy problem B

L'algorithme est à présent appliqué sur une corrélation de Green construite à partir de plusieurs temps de propagation. Dans cette configuration, nous allons montrer l'intérêt du cadre bayésien développé : nous parviendrons à estimer des temps de propagation non identifiables par l'application de l'identité de Ward. Sur cet exemple, une approche ad-hoc pour sélectionner le nombre de ruptures  $K$  est présentée, en alternative à l'utilisation du facteur de Bayes. Ce choix est motivé par la volonté de mettre en exergue des pathologies propres au choix manuel de  $K$ . L'estimation de la densité cible s'effectue en 2 passes, la première consistant à lancer une instance de l'algorithme d'estimation bayésienne avec un  $K$  surestimé par rapport à  $K_0$ . L'observation des densités estimées permettra de conclure quant à la pertinence de l'existence des ruptures associées. Pour ce faire, une mesure informative comme la variance des densités individuelles  $p(t_i|\mathbf{y}, K)$  peut être utilisée.

Les résultats de la première passe sont présentés dans la Fig. 4.7.

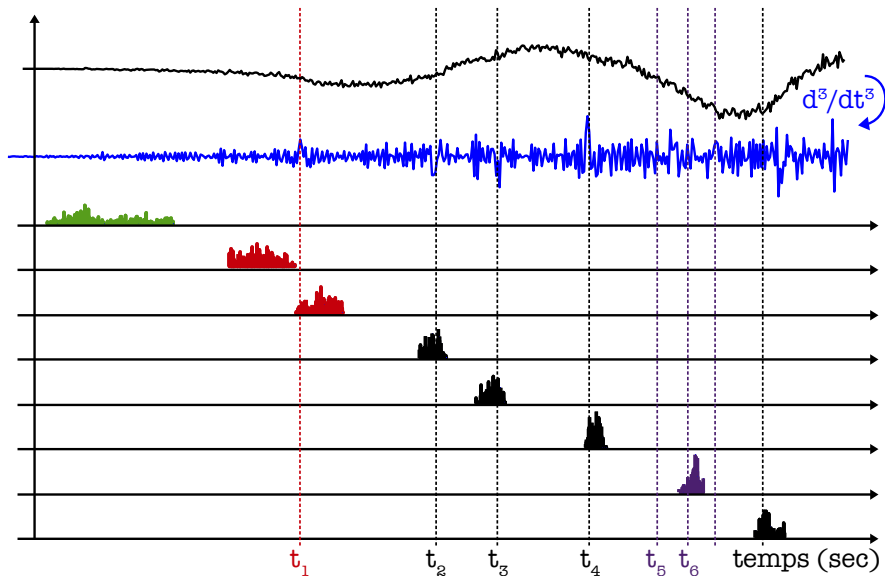


FIGURE 4.7 – Corrélation de Green et densités  $p(\mathbf{t}|\mathbf{y}, K = 8)$ .

Pour la première passe, huit densités sont estimées avec 1000 échantillons, chacune débutant à la première itération, de manière à couvrir uniformément le support de la fonctionnelle à régresser. Il apparaît clairement que certaines densités sont bien plus étalées que d'autres (variance 2 à 3 fois plus grande que la moyenne), ce qui conduit à douter de l'utilité/l'existence de la rupture. En particulier, la densité verte ne correspond à aucun temps de propagation et est la densité la plus étalée de toutes. Une deuxième pathologie, qui découle de ce que le nombre de ruptures recherché est sur-estimé, est que les densités rouge décrivent le même temps de propagation. Dans l'algorithme, une règle empêche les densités de se superposer pour éviter leur attraction commune vers la transition la plus évidente. Ici, les deux densités sont les plus proches l'une de l'autre possible et la décision peut être d'unifier les deux densités en diminuant le nombre de ruptures recherchées. Une troisième et dernière pathologie observée

est que la distribution violette est seule pour représenter 3 ruptures. C'est encore un résultat de ce que les densités ne peuvent pas se superposer. Par ailleurs, les temps sont très proches les uns des autres et la fonctionnelle ne présente pas de variations évidentes à cet endroit. Puisque le critère d'acceptation d'un mouvement est basé sur la fidélité de la régression, la contribution de transitions supplémentaires dans cette région est moins pertinente qu'ailleurs.

Les densités noires sont pertinentes et peuvent être conservées. Malgré la dégradation progressive du RSB avec le temps, il est notable que l'algorithme permet d'estimer les transitions sur l'intégralité du support de la fonctionnelle.

Dans l'exemple présenté, nous faisons le choix manuel de restreindre  $K$  à 6, pour se débarrasser de la densité verte et unifier les densités rouges. Une nouvelle instance de l'algorithme avec  $K = 6$  conduit aux résultats présentés dans la Fig. 4.8.

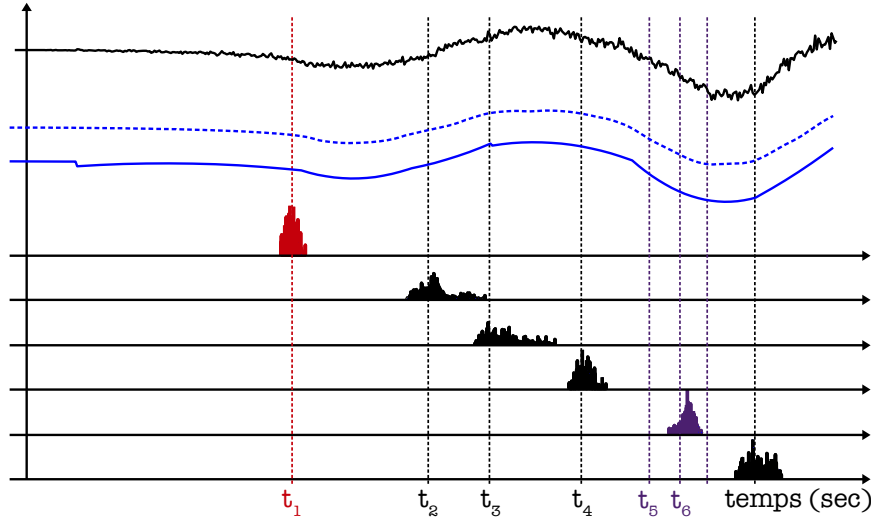


FIGURE 4.8 – Corrélation de Green et densités  $p(\mathbf{t}|\mathbf{y}, K = 8)$ .

Pour la deuxième passe, comme attendu, la fusion des densités rouges permet une très bonne identification du premier temps de propagation  $t_1$ , puisque c'est la densité de variance minimale parmi toutes celles estimées. La densité violette continue à jouer le rôle de 3 temps de propagation. En bleu figurent deux exemples de reconstruction de la corrélation de Green, par maximum de vraisemblance. La première, en pointillés, montre une reconstruction extrêmement fidèle correspondant au moyennage de 100 fonctionnelles reconstruite. La deuxième en trait plein correspond à l'une de celles estimées, dans un cas de « raté », dans le sens où la densité la plus à gauche s'est réfugiée sur la gauche du support temporel, ne correspondant plus à aucun temps de propagation. Ce cas illustre que malgré tout, la forme d'onde reconstruite est assez fidèle.

En conclusion, les toy problems présentés ont servis à caractériser la convergence de certaines statistiques des estimateurs déployés, dans un cas où l'identité de Ward ne permet pas de reconstruire correctement la forme d'onde la fonction de Green. Ces résultats préliminaires invitent à sur-estimer la période de chauffe, à défaut de maîtriser l'initialisation. Trois pa-

thologies liées au choix arbitraire du nombre de ruptures ont été identifiées, et l'approche proposée, construite sur une mesure de la variance des densités estimées, permet en deux passes, d'estimer une forme d'onde fidèle. D'ailleurs, une remarque conclusive importante est que la reconstruction des paramètres est sujet aux performances de régression. Dans notre cas, la fonctionnelle à régresser est de toute évidence assez lisse et ne présente pas de rupture évidente. Ainsi, le nombre de ruptures autant que leurs positions n'ont pas nécessairement besoin d'être très précis pour conduire à une reconstruction satisfaisante, expliquant les performances observées.

## 4.3 Raffinements algorithmiques et ouvertures

### 4.3.1 Evolution du nombre de morceaux

Dans l'algorithme présenté jusqu'à maintenant, le nombre  $K$  de morceaux de polynômes est fixé. Or, le choix de  $K$  est primordial dans le problème. Une possibilité pour le choisir est de considérer la densité de trajets  $D(t) = 4\pi c_0^3 \frac{t^2}{V}$  dans la réponse impulsionnelle qui permet alors de connaître, par intégration, le nombre de trajets lorsque la corrélation de Green est reconstruite sur une longueur donnée. Par exemple, pour une salle de dimensions  $4\text{m} \times 6\text{m} \times 3\text{m}$  en acoustique aérienne dans des conditions de pression et de température standard, 2 trajets sont attendus en dessous de 10 ms, 8 trajets en dessous de 15 ms, etc. Cette densité  $D(t)$  étant un résultat statistique moyen, le nombre  $K_0$  obtenu par intégration n'est pas toujours vrai et ne représente en réalité qu'un indicateur moyen.

Pour pallier cette incertitude, plusieurs modèles peuvent être mis en compétition : des instances de l'algorithme d'inférence bayésienne sont lancées pour différents modèles, c'est-à-dire pour des valeurs de  $K$  différentes, centrées autour de  $K_0$ . Pour juger de la pertinence d'un modèle plutôt que d'un autre, un score peut être attribué aux modèles. Un score possible est l'évidence du modèle : c'est-à-dire la vraisemblance des données marginalisées par rapport aux paramètres. En partant de la vraisemblance des paramètres :

$$p(\boldsymbol{\theta}|\mathbf{y}, K) = \frac{p(\mathbf{y}|\boldsymbol{\theta}, K)p(\boldsymbol{\theta}|K)}{p(\mathbf{y}|K)} \quad (4.28)$$

on peut écrire la log évidence :

$$\log p(\mathbf{y}|K) = \log p(\mathbf{y}|\boldsymbol{\theta}, K) - [\log p(\boldsymbol{\theta}|K) - \log p(\boldsymbol{\theta}|\mathbf{y}, K)] \quad (4.29)$$

qui contient la log vraisemblance  $\log p(\mathbf{y}|\boldsymbol{\theta}, K)$ , qui est en soi un bon critère pour quantifier la qualité d'un modèle. En revanche, ce terme n'intègre pas la notion de complexité du modèle, qui autorise logiquement un modèle complexe à avoir des performances au moins aussi bonnes qu'un modèle plus simple (en général). Se limiter à la vraisemblance pour comparer deux modèles donnerait la plupart du temps l'avantage au modèle le plus complexe. C'est là tout l'intérêt du second terme de la log évidence, qui va venir pénaliser cette complexité.

Pour comparer deux modèles entre eux, le facteur de Bayes [18]  $BF$  peut être utilisé. L'estimation du facteur de Bayes est difficile dans notre cas, en particulier à cause de la forme



complexe de l'évidence, où la marginalisation des paramètres n'est pas possible analytiquement. Certaines techniques d'estimation à partir de densité postérieure ont été étudiées, voir [19, 18].

Pour échapper à la difficulté d'estimation du facteur de Bayes, nous proposons à la place une transformation de l'algorithme MCMC vers un algorithme de type « *reversible-jump M.C.-M.C.* », c'est-à-dire comprenant des mouvements de chaîne consistant à changer le modèle, appelé mouvement de Green, [20, 21]. Ainsi, parce que ce genre de mouvement doit préserver la densité stationnaire de la chaîne et présenter un changement de modèle à tout instant réversible, 2 nouveaux mouvements sont introduits, en plus de la mise à jour d'une rupture :

1. Naissance d'un morceau supplémentaire
2. Mort d'un morceau

La Fig. 4.9 présente l'évolution du nombre de ruptures, incluant ces deux nouveaux mouvements. Dans celle-ci, les paramètres sont estimés pour différents modèles, la chaîne effectuant des sauts de l'un à l'autre. L'instant bleu est l'étape initiale, la chaîne étant créée pour estimer les paramètres de 5 morceaux. Le premier mouvement accepté par la chaîne est une réduction du modèle à 4 morceaux. A l'instant  $n$ , en rouge, 3 morceaux sont recherchés, mais ne sont plus les morceaux initiaux ; les temps des ruptures entre les morceaux ont été ré-initialisés à la création des morceaux. En pratique, il n'existe pas de rupture qui ne disparaissent jamais et la durée de vie d'une rupture est même plutôt courte (quelques dizaines de mouvements dans le toy-problem présenté), et ne permet par conséquent pas d'échantillonner suffisamment les densité marginales  $p(t_i)$ . Cette observation induit que cet algorithme permet l'estimation des densités jointes des paramètres, par modèle, où estimer la position du  $i$ ème temps de propagation en isolant la  $i$ ème composante de  $p(\mathbf{t}|K)$  n'a plus de sens.

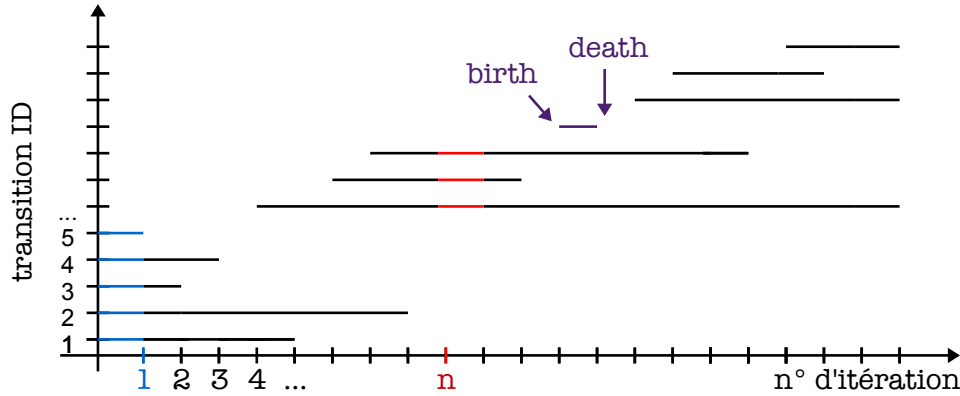


FIGURE 4.9 – Evolution typique des mouvements de naissance et de morts de transitions.

L'algorithme 4 résume l'algorithme « *reversible-jump M.C.-M.C.* ». Lorsque l'algorithme parcourt l'espace des modèles, la règle d'acceptation du mouvement ne rend pas chaque modèle aussi évident l'un que l'autre. Ainsi, si la complexité algorithmique est un verrou dans le déploiement de la méthode et qu'une sélection du modèle doit être opérée, le choix de  $K$  peut par exemple être réalisé par estimateur MAP à un certain moment de la chaîne :

$$\hat{K}_{MAP} = \arg \max_K p(K|\mathbf{y}) \quad (4.30)$$

**Algorithme 4:** Mouvements de Metropolis-Hastings-Green

---

<b>while</b> $n \leq N_{max}$ <b>do</b>	1
- choix équiprobable d'un mouvement parmi les trois	2
<b>if</b> $choix == naissance\ rupture$ <b>then</b>	3
- ordre du modèle augmenté : $K \leftarrow K + 1$	4
	5
<b>else</b>	6
- $choix == mort\ rupture$	7
- ordre du modèle diminué : $K \leftarrow K - 1$	8
<b>if</b> $choix == mise\ à\ jour\ d'une\ rupture$ <b>then</b>	9
- choix aléatoire $i$ d'une rupture	10
- mise à jour de la rupture $t_i$	11
- mise à jour des autres paramètres	12
<b>end</b>	13
- calcul du score obtenu avec la mise à jour	14
- test : mouvement accepté avec la probabilité $\alpha$	15
<b>if</b> $mouvement\ accepté$ <b>then</b>	16
- paramètres actuels conservés	17
- $n++$	18
<b>if</b> $n \geq n_{bi}$ <b>then</b>	19
- ajouter l'échantillon à la densité cible	20
<b>end</b>	21
<b>else</b>	22
- paramètres précédents restaurés	23
<b>end</b>	24
<b>end</b>	25

---

En procédant ainsi, l'algorithme RJMCMC redevient un problème sans saut dans la dimensionnalité du problème.

Nous proposons donc de découper l'algorithme complet en trois phases :

1. Période de chauffe ;
2. Sélection de modèle avec l'algorithme 4 « *reversible-jump M.C.-M.C.* » ;
3. Estimation des paramètres sur le modèle le plus vraisemblable par l'algorithme 3.

#### 4.3.1.1 $K$ et nouvelle densité cible

A présent, 3 mouvements composent la chaîne de Markov : l'ajout d'une rupture, la mort d'une rupture et la mise à jour des paramètres du modèle. Pour cette dernière étape, l'algorithme est identique à celui présenté précédemment, voir l'Algorithme 3, à ceci près qu'il est conditionnel au modèle en cours (c'est-à-dire à la valeur de  $K$ ). En revanche, pour les deux mouvements de Green, il convient maintenant de définir les probabilités d'accepter les mouvements.

Pour cela, présentons la nouvelle densité cible qui tient compte du modèle  $K$  étudié. Commençons par poser un *a priori* vague sur le nombre  $K$  de morceaux, à l'aide d'une loi de comptage telle qu'une loi de Poisson :

$$K|\lambda \sim Poi(\lambda), \text{ soit } p(K|\lambda) = \frac{\lambda^K \exp(-\lambda)}{K!} \quad (4.31)$$

où  $\lambda$  est un paramètre d'intensité, auquel est associé un *prior* peu informatif de Jeffrey,  $p(\lambda) = 1/\sqrt{\lambda} \mathbb{1}(\mathbb{R}_+)$ .

La marginalisation de  $\lambda$  s'écrit :

$$p(K) = \int_{\mathbb{R}} p(K|\lambda) p(\lambda) d\lambda = \frac{\int_{\mathbb{R}_+} \lambda^{K-\frac{1}{2}} \exp(-\lambda) d\lambda}{K!} = \frac{\Gamma(K + \frac{1}{2})}{\Gamma(K + 1)} \quad (4.32)$$

de telle sorte que :

$$p(\mathbf{t}, K) = p(\mathbf{t}|K) p(K) = \left(\frac{1}{L}\right)^K \times \frac{\Gamma(K + \frac{1}{2})}{\Gamma(K + 1)}. \quad (4.33)$$

Dans les Eq. (4.32) et (4.33),  $\Gamma(z + 1) = z!$  si  $z$  entier, sinon  $\int_{\mathbb{R}_+} t^{z-1} \exp(-t) dt$ .

Ainsi, la densité cible s'écrit maintenant :

$$p(\mathbf{t}, g|\mathbf{y}) \propto p(g) g^{-1} \times \left(\frac{1}{L}\right)^K \times \frac{\Gamma(K + \frac{1}{2})}{\Gamma(K + 1)} \times \frac{(1 + g)^{-3K/2}}{[\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_N - M\} \mathbf{y}]^{N/2}} \quad (4.34)$$

où

$$M = \frac{g}{1 + g} S [S^T \Lambda^{-1} S]^{-1} S^T \Lambda^{-1}. \quad (4.35)$$

#### 4.3.1.2 Mouvements de Green

Dans le problème exposé, deux mouvements équiprobables (1/3) changent la dimension du modèle. Dans ses travaux, Green définit la probabilité d'accepter un mouvement de la chaîne qui intègre ce fait :

$$r_m = \frac{p(\tilde{\mathbf{t}}, \tilde{g}|\mathbf{y})}{p(\mathbf{t}, g|\mathbf{y})} \times \frac{p_m(\tilde{\mathbf{t}}, \tilde{g})}{p_m(\mathbf{t}, g)} \times \frac{q_m(\tilde{u})}{q_m(u)} \times J(\text{mouvement}) \quad (4.36)$$

où les variables tildées sont celles mises à jour avec un changement de dimension,  $p_m$  est la probabilité de tenter le mouvement  $m$  (-1 pour une mort, +1 pour une naissance) et  $q_m$  est la loi associée à la génération de la variable auxiliaire  $u$  du mouvement. La Jacobienne  $J$  qui rend compte du changement de variable et de dimension est égale à l'unité dans notre cas, car les variables auxiliaires sont indépendantes des paramètres. Pour illustrer la notion de variable auxiliaire, considérons  $u \sim \mathcal{N}(0, \sigma^2)$  et  $x' = x + u$ . Nous avons  $u' = -u$ , variable auxiliaire du mouvement inverse :

$$t : (x, u) \mapsto (x + u, -u) \quad (4.37)$$

alors :  $J_t = \left| \frac{\partial(x+u, -u)}{\partial(x, u)} \right| = 1$ .

Pour un mouvement de naissance, ce ratio s'écrit alors :

$$r_{+1} = \frac{p(\tilde{\mathbf{t}}, \tilde{g}|\mathbf{y})}{p(\mathbf{t}, g|\mathbf{y})} \times \frac{p_{-1}(\tilde{\mathbf{t}}, \tilde{g})}{p_{+1}(\mathbf{t}, g)} \times \frac{1}{p(\tilde{t}, \tilde{K})} \quad (4.38)$$

où

$$\frac{p(\tilde{\mathbf{t}}, \tilde{g}|\mathbf{y})}{p(\mathbf{t}, g|\mathbf{y})} = \frac{(1 + \tilde{g})^{3(K+2)/2}}{(1 + g)^{3(K+1)/2}} \times \left( \frac{\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_N - M\} \mathbf{y}}{\mathbf{y}^T \tilde{\Lambda}^{-1} \{\mathbb{1}_N - \tilde{M}\} \mathbf{y}} \right)^{\frac{N}{2}} \times \frac{K + \frac{1}{2}}{(K + 1)L} \quad (4.39)$$

en utilisant le fait que  $\Gamma(K + 1) = K\Gamma(K)$ . Par ailleurs, on sait également que  $p_{-1}(\tilde{\mathbf{t}}, \tilde{g}) = 1/(K + 1)$ ,  $p_{+1}(\mathbf{t}, g) = 1/L$  et  $p(\tilde{t}, \tilde{K}) = 1/L$ , donc :

$$r_{+1} = \frac{(1 + \tilde{g})^{3(K+2)/2}}{(1 + g)^{3(K+1)/2}} \times \left( \frac{\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_N - M\} \mathbf{y}}{\mathbf{y}^T \tilde{\Lambda}^{-1} \{\mathbb{1}_N - \tilde{M}\} \mathbf{y}} \right)^{\frac{N}{2}} \times \frac{K + \frac{1}{2}}{(K + 1)^2} \quad (4.40)$$

Pour un mouvement de mort, le ratio s'écrit :

$$r_{-1} = \frac{p(\tilde{\mathbf{t}}, \tilde{g}|\mathbf{y})}{p(\mathbf{t}, g|\mathbf{y})} \times \frac{p_{+1}(\tilde{\mathbf{t}}, \tilde{g})}{p_{-1}(\mathbf{t}, g)} \times p(t, K) \quad (4.41)$$

où

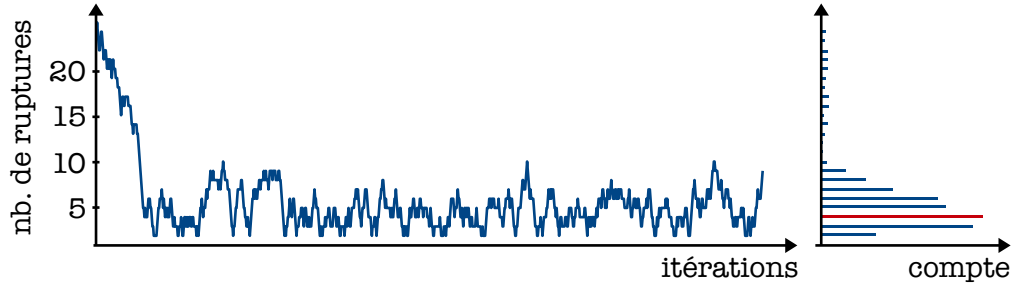
$$\frac{p(\tilde{\mathbf{t}}, \tilde{g}|\mathbf{y})}{p(\mathbf{t}, g|\mathbf{y})} = \frac{(1 + \tilde{g})^{3K/2}}{(1 + g)^{3(K+1)/2}} \times \left( \frac{\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_N - M\} \mathbf{y}}{\mathbf{y}^T \tilde{\Lambda}^{-1} \{\mathbb{1}_N - \tilde{M}\} \mathbf{y}} \right)^{\frac{N}{2}} \times \frac{KL}{K - \frac{1}{2}} \quad (4.42)$$

en utilisant le fait que  $\Gamma(K + 1) = K\Gamma(K)$ . Par ailleurs, on sait également que  $p_{+1}(\tilde{\mathbf{t}}, \tilde{g}) = 1/L$ ,  $p_{-1}(\mathbf{t}, g) = 1/K$  et  $p(t, K) = 1/L$ , donc :

$$r_{-1} = \frac{(1 + \tilde{g})^{3K/2}}{(1 + g)^{3(K+1)/2}} \times \left( \frac{\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_N - M\} \mathbf{y}}{\mathbf{y}^T \tilde{\Lambda}^{-1} \{\mathbb{1}_N - \tilde{M}\} \mathbf{y}} \right)^{\frac{N}{2}} \times \frac{K^2}{K - \frac{1}{2}}. \quad (4.43)$$

#### 4.3.1.3 Application à la corrélation de Green

Dans la Fig. 4.10, nous illustrons l'étape de sélection de modèle. Les deux autres étapes (chauffe et estimation des paramètres) sont identiques à l'algorithme 3. Initialement, pour illustrer la convergence de la chaîne, un très grand nombre de ruptures sont recherchées. En quelques dizaines d'itérations il devient possible de commencer l'échantillonnage de la densité postérieure de  $K$ . Quelques centaines de mouvements plus tard, l'estimateur MAP est appliqué pour retenir le nombre de ruptures le plus vraisemblable.

FIGURE 4.10 – Choix du nombre de ruptures  $K$  par maximum *a posteriori*.

### 4.3.2 Filtrage des données

Une des conclusions fortes des sections précédentes est que le degré de modélisation présenté jusqu'à présent n'est pas suffisant pour rendre compte de la corrélation de Green estimée. Son modèle théorique, construit en combinant un modèle à retards pour la fonction de Green avec l'identité de Ward est bien polynomial d'ordre 2 par morceaux. En revanche, ce modèle n'intègre pas l'estimation par la mesure de signaux réels et qui peut se traduire par un filtrage (notamment le caractère bande passante et non blanc des signaux traités). Afin de prendre en compte ce filtrage dans l'algorithme d'estimation, plusieurs ouvertures sont proposées :

1. Modéliser la chaîne d'acquisition par un FIR  $h$ . Le filtre correspondant au spectre des données utilisées étant identifié, nous proposons de l'imposer au modèle d'observation, de telle sorte qu'il devienne

$$\mathbf{y} \leftarrow H\mathbf{S}\boldsymbol{\beta} + \mathbf{w} \quad (4.44)$$

où  $H$  est la matrice de convolution du filtre  $h$ . Ainsi, les paramètres à estimer sont toujours les paramètres originaux, correspondant au modèle théorique et juste de la corrélation de Green, mais la vraisemblance du modèle change, puisqu'elle tient maintenant compte du filtrage. Le bruit d'estimation  $\mathbf{w}$  voit son modèle inchangé puisqu'il reste un résidu d'estimation de corrélation, mais les valeurs des paramètres qui le décrivent ne sont vraisemblablement plus les mêmes.

2. Le filtre  $h$  peut être inversé pour créer de nouvelles données et ainsi ramener l'observation au niveau du modèle théorique de polynômes par morceaux :

$$\mathbf{y} \leftarrow \mathbf{h}^{-1} * \mathbf{y} \quad (4.45)$$

en employant par exemple un filtrage de Kirkeby soigneusement calculé.

3. Cependant, l'inversion d'un filtre est sujet à des problèmes d'instabilité. Une autre approche consiste à considérer la combinaison du filtrage des données et du modèle polynomial, pour être approchée par un polynôme. L'ordre de ce polynôme, logiquement plus grand que 2, n'aurait eu d'impact sur l'algorithme qu'un changement de dimension :

$$\mathbf{y} = \mathbf{S}\boldsymbol{\beta} + \mathbf{w}, \boldsymbol{\beta} \in \mathbb{R}^{\geq 3(K+1)} \quad (4.46)$$

En revanche, les paramètres  $\mathbf{t}$  et  $\boldsymbol{\beta}$  ne seraient plus liés au modèle théorique de la corrélation de Green et ce faisant, il ne serait plus possible d'estimer le moindre paramètre physique.

L'unique intérêt aurait résidé dans l'estimation de la forme de la corrélation de Green, condensée dans un jeu de paramètres peu nombreux.

#### 4.3.3 Conclusion

Dans ce dernier chapitre, un nouveau modèle pour la corrélation de Green a été proposé et un cadre d'estimation bayésienne a été construit autour d'un algorithme de type Chaîne de Markov Monte-Carlo, finalement étendu pour intégrer le choix de la dimensionnalité du modèle ( $K$ ) dans les mouvements de la chaîne. Les estimateurs par *maximum a posteriori* des paramètres transforme l'estimation des paramètres polynomiaux en un problème de régression sous contrainte. Cette approche a permis l'estimation des paramètres tels que les temps de propagation et les amplitudes associées dans des cas où l'identité de Ward véhiculait une propagation d'erreur trop grande pour être exploitée. La dernière étape, pour pouvoir appliquer l'algorithme sur des données réelles, consiste à tenir compte du fait que la corrélation de Green est en pratique estimée par la corrélation de codas acoustiques.



# Références

- [1] M. SANQUER, F. CHATELAIN, M. EL-GUEDRI et N. MARTIN, “Modèle polynomial par morceaux muni de transitions régulières - Application à la modélisation de signaux transitoires électriques”, dans *GRETSI*, 2011, p. 1–4 (cf. p. 86, 89, 90).
- [2] —, “A smooth transition model for multiple-regime time series”, *IEEE Transactions on signal processing*, t. 61, n° 7, p. 3960–3963, 2013 (cf. p. 86, 89, 90).
- [3] E. T. JAYNES, “Prior Probabilities”, *IEEE Transactions on System, Science and Cybernetics*, t. 4, n° 3, p. 227–241, 1968 (cf. p. 89).
- [4] D. J. C. MACKAY, “Bayesian Interpolation”, *Neural Computation*, t. 4, p. 415–447, 1992 (cf. p. 89).
- [5] —, “Comparison of approximate methods for handling hyperparameters”, *Neural Computation*, t. 11, n° 5, 1999 (cf. p. 89, 90).
- [6] B. K. M. ALLICK et R. J. C. ARROLL, “Spatially Adaptive Bayesian Penalized Regression Splines ( P-splines )”, *Journal of Computational and Graphical Statistics*, t. 14, n° 2, p. 378–394, 2005 (cf. p. 90).
- [7] E. PUNSKAYA, C. ANDRIEU, A. DOUCET et W. J. FITZGERALD, “Bayesian Curve Fitting Using MCMC With Applications to Signal Segmentation”, *IEEE Transactions on signal processing*, t. 50, n° 3, p. 747–758, 2002 (cf. p. 90).
- [8] N. DOBIGEON, J.-y. TOURNERET et M. DAVY, “Joint Segmentation of Piecewise Constant Autoregressive Processes by Using a Hierarchical Model and a Bayesian Sampling Approach”, *IEEE Transactions on signal processing*, t. 55, n° 4, p. 1251–1263, 2007 (cf. p. 90).
- [9] D. G. T. DENISON, B. K. MALLICK et A. F. M. SMITH, “Automatic Bayesian curve fitting”, *Journal of the Royal Statistical Society*, t. 60, n° 2, p. 333–350, 2007 (cf. p. 90).
- [10] H. JEFFREYS, “An invariant form for the prior probability in estimation problems”, *Proceedings of the Royal Society of London*, t. 186, n° 1007, p. 453–461, 1946 (cf. p. 90).
- [11] B. LISEO, “Elimination of nuisance parameters with reference noninformative priors”, rap. tech., 1990 (cf. p. 90).
- [12] W. K. HASTINGS, “Monte Carlo Sampling Methods Using Markov Chains and Their Applications”, *Biometrika*, t. 57, n° 1, p. 97–109, 1970 (cf. p. 91).
- [13] P. H. PESKUN, “Optimum Monte-Carlo sampling using Markov chains”, *Biometrika*, t. 60, n° 3, p. 607–612, 1973 (cf. p. 92).
- [14] T. PARK et D. VAN DYCK, “Partially Collapsed Gibbs Samplers : Illustrations and Applications”, *American Statistical Association*, t. 18, n° 2, p. 283–305, 2009 (cf. p. 93).
- [15] S. ARLOT et A. CELISSE, “A survey of cross-validation procedures for model selection”, *Statistics Surveys*, t. 4, p. 40–79, 2010 (cf. p. 96).
- [16] P. C. HANSEN, “The L-curve and its use in the numerical treatment of inverse problems”, rap. tech., p. 1–24 (cf. p. 96).



- 
- [17] R. VINCENT, “Master Internship Report : Regression with kernels : application to Granger causality and online adaptative optimisation algorithms”, 2012 (cf. p. 96).
  - [18] J. M. SATAGOPAN, M. A. NEWTON et A. E. RAFTERY, “Easy estimation of normalizing constants and Bayes factors from posterior simulation : Stabilizing the harmonic mean estimator”, rap. tech., 2000 (cf. p. 101, 102).
  - [19] S. M. LEWIS et A. E. RAFTERY, “Estimating Bayes Factors via Posterior Simulation with the Laplace-Metropolis Estimator”, *Journal of the American Statistical Association*, t. 92, n° 438, 1997 (cf. p. 102).
  - [20] P. J. GREEN, “Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination”, *Biometrika*, t. 82, n° 4, p. 711–732, 1995 (cf. p. 102).
  - [21] P. J. GREEN, “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination”, *Biometrika*, t. 82, n° 4, p. 711–732, 1995 (cf. p. 102).

# Conclusions et perspectives

Ainsi s’achève ce manuscrit, sur une note récapitulative et d’ouverture. C’est une note optimiste, car écrite après avoir travaillé trois ans sur un sujet passionnant aux côtés d’amis extraordinaires par leurs compétences et leur implication. Des perspectives, il y en a bien sûr, car comme à l’accoutumée, le nombre de réponses apportées n’égale pas celui des questions soulevées en chemin. Mais c’est là, je trouve, que réside toute la beauté de la recherche.

Deux axes de recherche ont été formulés pour traiter la thématique : « estimation dans un contexte passif ». Le premier axe consiste à caractériser les performances d’estimateurs dans un contexte passif d’estimation, en regard du contenu spectral temporel et spatial des champs mesurés et de leur modélisation. Le deuxième axe permet de diversifier les applications des méthodes d’estimation passives en se concentrant sur une thématique à enjeux : le Structural Health Monitoring.

## Résumé technique

Capitalisant sur les travaux de doctorat de M. Carmona [1], c’est une notion de filtrage élémentaire et puissante qui a constitué notre point d’entrée sur la problématique : la formule des interférences. Etendue à un milieu de propagation visco-acoustique, elle a permis d’obtenir des modèles d’observations pour le contexte passif d’estimation. Ces relations, fondamentales pour étudier des estimateurs et formaliser leurs performances théoriques, parviennent à rendre compte des propriétés spectrales temporelles et/ou spatiales des champs mesurés. Divers estimateurs sont construits : corrélation de Green, fonction de Green, temps de propagation. . . sur des modèles d’observation qui s’enchevêtrent et où les estimateurs propagent les incertitudes des modèles amonts. Leurs propriétés de précision, de vitesse de convergence et de robustesse ont été illustrées, par la simulation et par l’expérimentation dans des milieux à géométrie contrôlée.

Une ouverture applicative a été proposée, en portant la modalité acoustique audible aérienne au monde du SHM. Elle a été réalisée en proposant l’étude de sensibilité des estimateurs aux variations de la géométrie du milieu et des exemples de paramètres reconstruits pour l’occasion sont la densité spectrale de puissance, le module d’Young d’une poutre ou encore la géométrie d’un réseau de capteurs. A également été proposée la description fonctionnelle d’un noeud de mesure acoustique prodiguant le calcul des estimateurs en contexte passif et en embarqué. Pour adapter les algorithmes, nous avons exploité la finalité qui consiste à calculer des corrélations de champs gaussiens, afin de réduire la quantité de données à traiter. La quantification des champs a été reconsidérée pour finalement ne garder que leur signe, couplée à l’extraction analogique des champs réverbérés et à leur échantillonnage dans une bande de fréquences adaptée.

S’il n’y avait qu’une conclusion à retenir de tous ces travaux, c’est que la corrélation de

bruit ambiant, qui fournit une image de la corrélation de Green, voit son estimation être la plus robuste et la plus rapide de tous les estimateurs étudiés. A la lumière de ce résultat, un ultime modèle a été élaboré, pour la corrélation de Green, tendant la perche aux méthodes d'inférences bayésiennes, qui offrent l'accès à divers paramètres contenus dans la fonction de Green sans avoir à estimer cette dernière à proprement parler. Des estimateurs par maximum *a posteriori* sur les densités estimées par un algorithme de type MCMC ont été développés, implémentés et validés.

## Compléments de recherche

De nombreux raffinements des travaux proposés peuvent être menés en regard d'une application et d'un besoin spécifiques. Cependant, deux perspectives générales importantes se dégagent et mériteraient d'être approfondies.

### A propos du degré de modélisation ?

Le choix du niveau de modélisation de la physique influe directement sur la fonctionnalité des estimateurs et leurs performances. En ce sens, le modèle quadratique par morceaux adopté pour la corrélation de codas est incomplet, avec pour preuve le désaccord modéré entre les résultats théoriques et expérimentaux. Ce constat ne met cependant pas un point final à cette ramification de ma recherche, car la corrélation de Green est une quantité fondamentale dans la sphère de l'estimation en contexte passif, et je vois donc un intérêt à s'obstiner. Un degré de modélisation supplémentaire, qui permettrait d'améliorer les performances des estimateurs, serait de mieux rendre compte du contenu spectral des champs mesurés et pour cela, le savoir-faire en acoustique des salles constituerait un point de départ sérieux. Je suggère pour ce faire l'égalisation de la chaîne d'acquisition ainsi que l'étude de l'effet de la directionnalité du capteur. Une méthode envisagée serait une version adaptative du filtrage de Kikerby, filtrage adapté de la réponse impulsionnelle de la chaîne de mesure où une contrainte adaptative sur la bande passante du champ permet de tenir compte de la décroissance du RSB due à l'atténuation d'amplitude du champ par propagation. L'algorithme de régression polynomiale par morceaux de la corrélation de Green bénéficierait du filtrage de Kikerby et permettrait de faire un pas vers l'applicabilité de la méthode sur des données réelles.

Le cadre bayésien présenté pour la régression de la corrélation de Green est pertinent mais primitif. Il est légitime car il offre une réponse naturelle à la problématique de localisation de ruptures dans un modèle continu par morceaux. Cependant, en regard de l'état de l'art, l'algorithme utilisé ne représente un incrément qu'au niveau du modèle d'observation (bruit hétéroscédastique, matrice de design inhabituelle). Une amélioration serait de rendre l'algorithme capable de sélectionner automatiquement le nombre de ruptures présent dans la fenêtre d'observation de travail. Pour ce faire, le facteur de Bayes est un outil qui peut être supplanté par un algorithme de type *reversible-jump Markov chain Monte Carlo*. Dans celui-ci, la collection des mouvements de la chaîne possibles se voit agrémentée de l'augmentation/réduction

du nombre de ruptures cherchées dans le modèle. Ce genre de mouvement, qui consiste à changer la dimensionnalité du modèle, appartient à la classe des mouvements de Green et est déjà très étudiée dans la littérature. L'algorithme a d'ores et déjà été implémenté et un article de journal présentant l'extension des travaux du manuscrit sur l'estimation bayésienne est en cours de rédaction.

Enfin, j'aimerais insister sur le fait que la dissipation est le point clef pour étendre l'approche d'estimation en contexte passif à d'autres modalités. Ma recommandation à ce sujet est très simple : c'est le point à traiter en premier, puisqu'il conditionne l'identité de Ward du milieu étudié. C'est également le point le plus délicat car pour être exploitable et interprétable, l'identité peut se transformer en approximation, auquel cas l'erreur commise mérite d'être étudiée. Je parle d'extension à d'autres modalités que l'acoustique aérienne audible, car j'identifie au moins deux pistes qui peuvent présenter un intérêt applicatif dans le monde du SHM ou des télécommunications : la propagation acoustique enfouie, par exemple dans du béton, et la propagation électro-magnétique, par exemple RF. Dans le premier cas, l'idée serait d'identifier la structure depuis son intérieur et non son extérieur comme c'était le cas en acoustique aérienne, moyennant de nouveaux modèles sur la dissipation et la vitesse de propagation. Dans le deuxième cas, l'approche passive pourrait intervenir pour l'estimation de canal ou encore l'auto-localisation d'un réseau GSM par exemple.

### A propos de la coda ?

Dans les travaux exposés, le contexte passif d'estimation a été associé à l'écoute de sources d'opportunité, et plus exactement de leurs champs réverbérés, ayant acquis de l'information sur le milieu en se propageant. Ce que nous savons, c'est qu'afin d'appliquer la formule des interférences, la coda est supposée localement stationnaire et les résultats expérimentaux tendent à valider cette hypothèse. La coda est également un champ vraisemblablement gaussien centré ; observée comme telle expérimentalement, et peut se justifier en invoquant le théorème central-limite sur la densité de trajets en interférence. Si le milieu de propagation est linéaire, le contenu spectral du champ mesuré est à l'image de la source qui l'a généré. Au cours de la propagation, le champ subit une atténuation géométrique et par dissipation qui conduit à la réduction de la bande fréquentielle temporelle apparente. Enfin, la coda acoustique est fortement corrélée dans l'espace, puisqu'elle conduit expérimentalement à la corrélation de Green du milieu. Pour expliquer pourquoi, les notions d'équipartition, de chaos et d'ergodicité, longuement discutées dans la littérature, forment le point d'entrée vers ce sujet d'ouverture intéressant mais vraisemblablement abyssal.

Finalement, j'attire l'attention du lecteur, qui en a maintenant terminé avec le manuscrit, sur la thématique du calcul de fonctions de corrélation dans un espace compressé, de champs normalisés et/ou fortement quantifiés. Le calcul de corrélations de champs gaussiens à partir de leur signe, où l'information sur l'énergie du signal est perdue, est le point abordé dans le manuscrit. Resurgissant pour répondre à la problématique de l'embarqué et offrir une solution aux contraintes de limitations sur les quantités de données manipulées, ces travaux des années 60 n'ont pas encore dévoilé tout leur potentiel. Je souhaite en effet proposer une ouverture sur

le sujet, en utilisant un algorithme de calcul de corrélation de complexité très réduite, puisque ne faisant intervenir que des sommes : la méthode du décréement aléatoire, empruntée aux techniques du SHM, elles aussi datant des années 60. Une introduction au sujet est proposée dans l'Annexe E. Ce faisant, c'est la notion de filtrage qui pourrait être intégralement revue pour le calcul embarqué, travaillant sur des signaux quantifiés radicalement et moyennant une complexité plus faible que les méthodes usuelles.

## Référence

- [1] M. CARMONA, "Identification passive des milieux. Application à la reconstruction de géométries dans les réseaux de capteurs et au diagnostic des structures", thèse de doct., Université de Grenoble, 2008 (cf. p. 111).

# Annexes de l'introduction

## A.1 L'oscillateur amorti

Le système considéré peut être étudié sous forme de circuit électrique :

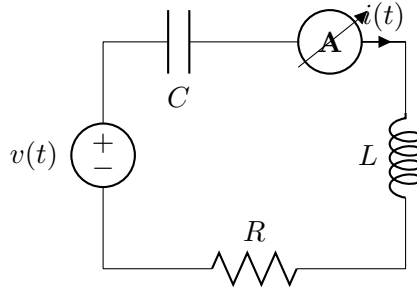


FIGURE A.1 – Circuit RLC, oscillateur amorti

La charge de courant  $q(t)$  obéit à une équation différentielle du second degré en temps, qui s'obtient par exemple en considérant la loi des mailles sur l'unique boucle du circuit :

$$\ddot{q}(t) + 2\eta\dot{q}(t) + \omega_0^2 q(t) = \frac{1}{L}v(t) \quad (\text{A.1})$$

où  $\omega_0 = \sqrt{1/LC}$  est la pulsation propre du circuit à laquelle la charge devrait osciller en l'absence d'amortissement, ce dernier étant modélisé par la constante  $\eta = R/2L$  (pertes par effet Joule).

La réponse de l'oscillateur à une impulsion est sa réponse à la sollicitation la plus élémentaire qui soit et suffit à décrire exhaustivement le système. Pour l'oscillateur amorti, elle est notée  $h(t)$  et peut se déduire de l'équation différentielle (A.1) en passant dans le domaine de Laplace. En prenant pour condition initiale le condensateur déchargé,  $h(t)$  est de la forme :

$$h(t) = \frac{1}{L} \frac{1}{\omega} \exp(-\eta t) \sin(\omega t) u(t). \quad (\text{A.2})$$

**Remarque :** les conclusions obtenues dans la suite des calculs pourraient également être obtenues en considérant un condensateur initialement chargé (et donc une solution en  $\cos()$  et  $\sin()$ ). Le coefficient multiplicatif  $1/L$  ne change pas non plus la nature des conclusions. Nous verrons qu'il introduira simplement une mise à l'échelle du résultat.

$u(t)$  est l'échelon de Heaviside, tenant compte du fait que  $h(t)$  est une fonction causale. L'amortissement se traduit par le terme monotone décroissant  $\exp(-\eta t)$  et se retrouve également dans la partie oscillante de la réponse, en modifiant la pulsation d'oscillation  $\omega_0$  du système non amorti par une pulsation plus faible  $\omega = \sqrt{\omega_0^2 - \eta^2}$ .

## A.2 Corrélation de Green

L'objectif est d'exhiber un lien fonctionnel qui relie la réponse impulsionnelle  $h(t)$  à sa fonction d'auto-corrélation  $C_h(t)$ . Définissons donc la fonction de corrélation d'une fonction  $f(t)$  réelle monovariée stationnaire :

$$C_f(t) \triangleq \int_{\mathbb{R}} f(t') f(t' + t) dt'. \quad (\text{A.3})$$

La corrélation de  $h(t)$  pour l'oscillateur amorti s'écrit :

$$C_h(t) = \frac{1}{4L^2\omega_0^2} \exp(-\eta|t|) \left[ \frac{\cos(\omega t)}{\eta} + \frac{\sin(\omega|t|)}{\omega} \right]. \quad (\text{A.4})$$

*Démonstration.* En partant de la définition de la corrélation (A.4), nous avons :

$$C_h(\tau) = \frac{1}{L^2} \frac{1}{\omega^2} \int_{\mathbb{R}} u(t) u(t + \tau) \exp(-\eta t) \sin(\omega t) \exp(-\eta(t + \tau)) \sin(\omega(t + \tau)) dt \quad (\text{A.5})$$

où les échelons de Heaviside ont un impact sur le support d'intégration :

$$C_h(\tau) = \frac{1}{L^2} \frac{1}{\omega^2} \exp(-\eta\tau) \int_{\min(0, -\tau)}^{+\infty} \exp(-2\eta t) \sin(\omega t) \sin(\omega(t + \tau)) dt. \quad (\text{A.6})$$

Sachant que  $2\sin(a)\sin(b) = \cos(a - b) - \cos(a + b)$ , nous obtenons :

$$C_h(t) = \frac{1}{2L^2} \frac{1}{\omega^2} \exp(-\eta\tau) (I_1 - I_2) \quad (\text{A.7})$$

avec :

$$I_1 = \cos(\omega\tau) \int_{\min(0, -\tau)}^{+\infty} \exp(-2\eta t) dt = \frac{1}{2\eta} \cos(\omega\tau) \exp(-2\eta \min(0, -\tau)) \quad (\text{A.8})$$

et  $I_2$  se calcule à l'aide de deux intégrations par parties successives. En effet,

$$I_2 = \int_{\min(0, -\tau)}^{+\infty} \exp(-2\eta t) \cos(2\omega t + \omega\tau) dt \quad (\text{A.9})$$

et en posant  $f(t) = \cos(2\omega t + \omega\tau)$  et  $g'(t) = \exp(-2\eta t)$ , nous obtenons :

$$I_2 = \left[ \cos(2\omega t + \omega\tau) \frac{\exp(-2\eta t)}{-2\eta} \right]_{+\infty}^{\min(0, -\tau)} - \int_{\min(0, -\tau)}^{+\infty} \frac{-2\omega}{-2\eta} \exp(-2\eta t) \sin(2\omega t + \omega\tau) dt$$

$$(A.10)$$

ce qui s'écrit encore :

$$I_2 = \frac{1}{2\eta} \cos(2\omega|\tau|) \exp(-2\eta \min(0, -\tau)) - \frac{\omega}{\eta} \int_{\min(0, -\tau)}^{+\infty} \exp(-2\eta t) \sin(2\omega t + \omega\tau) dt \quad (A.11)$$

et qui se simplifie encore en :

$$I_2 = \frac{1}{2\eta} \cos(\omega\tau) \exp(-2\eta \min(0, -\tau)) - \frac{\omega}{\eta} I_3 \quad (A.12)$$

où

$$I_3 = \int_{\min(0, -\tau)}^{+\infty} \exp(-2\eta t) \sin(2\omega t + \omega\tau) dt. \quad (A.13)$$

$I_3$  se calcule à nouveau par intégration par parties, pour faire apparaître  $I_2$ . En posant  $f(t) = \sin(2\omega t + \omega\tau)$  et  $g'(t) = \exp(-2\eta t)$ , nous obtenons :

$$I_3 = \left[ \sin(2\omega t + \omega\tau) \frac{\exp(-2\eta t)}{-2\eta} \right]_{+\infty}^{\min(0, -\tau)} - \int_{\min(0, -\tau)}^{+\infty} \frac{2\omega}{-2\eta} \exp(-2\eta t) \cos(2\omega t + \omega\tau) dt \quad (A.14)$$

ce qui s'écrit encore :

$$I_3 = \frac{1}{2\eta} \sin(\omega|\tau|) \exp(-2\eta \min(0, -\tau)) + \frac{\omega}{\eta} I_2 \quad (A.15)$$

Ce faisant, nous pouvons maintenant écrire l'expression de  $I_2$  :

$$I_2 = \frac{1}{2\eta} \frac{\eta^2}{\eta^2 + \omega^2} \exp(-2\eta \min(0, -\tau)) \left[ \cos(\omega\tau) - \frac{\omega}{\eta} \sin(\omega|\tau|) \right] \quad (A.16)$$

En regroupant tous ces résultats intermédiaires, nous obtenons :

$$I_1 - I_2 = \frac{1}{2} \frac{1}{\eta} \exp(-2\eta \min(0, -\tau)) \left[ \cos(\omega\tau) - \frac{\eta^2}{\eta^2 + \omega^2} \left[ \cos(\omega\tau) + \frac{\omega}{\eta} \sin(\omega|\tau|) \right] \right] \quad (A.17)$$

et donc par conséquent :

$$C_h(t) = \frac{1}{4L^2} \exp(-\eta|\tau|) \left[ \frac{1}{\eta} \frac{1}{\eta^2 + \omega^2} \cos(\omega\tau) + \frac{1}{\omega} \frac{1}{\eta^2 + \omega^2} \sin(\omega|\tau|) \right], \quad (A.18)$$

soit finalement le résultat annoncé :

$$C_h(t) = \frac{1}{4} \frac{1}{L^2} \frac{1}{\omega_0^2} \exp(-\eta|\tau|) \left[ \frac{1}{\eta} \cos(\omega\tau) + \frac{1}{\omega} \sin(\omega|\tau|) \right]. \quad (A.19)$$

□



### A.2.1 Dérivée de la corrélation de Green

La dérivée première de  $C_h(t)$  s'écrit :

$$\dot{C}_h(t) = -\frac{1}{4} \frac{1}{L^2} \frac{1}{\omega_0^2} \exp(-\eta|t|) \left[ \frac{\eta}{\omega} + \frac{\omega}{\eta} \right] \sin(\omega t). \quad (\text{A.20})$$

*Démonstration.* Calculons

$$D_1 = \frac{d}{dt} \exp(-\eta|\tau|) \left[ \frac{1}{\eta} \cos(\omega\tau) + \frac{1}{\omega} \sin(\omega|\tau|) \right]. \quad (\text{A.21})$$

Nous avons :

$$D_1 = -\eta \operatorname{sgn}(\tau) \exp(-\eta|\tau|) \left[ \frac{1}{\eta} \cos(\omega\tau) + \frac{1}{\omega} \sin(\omega|\tau|) \right] + \exp(-\eta|\tau|) \left[ -\frac{\omega}{\eta} \sin(\omega\tau) + \cos(\omega\tau) \right] \quad (\text{A.22})$$

ce qui permet d'obtenir directement le résultat :

$$D_1 = -\exp(-\eta|t|) \left[ \frac{\eta}{\omega} \operatorname{sgn}(\tau) \sin(\omega|t|) + \frac{\omega}{\eta} \sin(\omega t) \right] = -\exp(-\eta|t|) \left[ \frac{\eta}{\omega} + \frac{\omega}{\eta} \right] \sin(\omega t). \quad (\text{A.23})$$

□

## A.3 Identité de Ward

Nous cherchons le lien fonctionnel entre :

$$h(t) = \frac{1}{L} \frac{1}{\omega} \exp(-\eta t) \sin(\omega t) u(t) \quad (\text{A.24})$$

et

$$\dot{C}_h(t) = -\frac{1}{4} \frac{1}{L^2} \frac{1}{\omega_0^2} \exp(-\eta|t|) \left[ \frac{\eta}{\omega} + \frac{\omega}{\eta} \right] \sin(\omega t). \quad (\text{A.25})$$

L'identité de Ward pour l'oscillateur amorti s'écrit :

$$h(t) = -4\eta L \dot{C}_h(t) u(t). \quad (\text{A.26})$$

*Démonstration.* Il suffit d'observer que :

$$\frac{\eta}{\omega} + \frac{\omega}{\eta} = \frac{\omega_0^2}{\omega\eta} \quad \rightarrow \quad \dot{C}_h(t) = -\frac{1}{4} \frac{1}{L^2} \exp(-\eta|t|) \frac{1}{\omega\eta} \sin(\omega t). \quad (\text{A.27})$$

□

## A.4 Extension

A présent, nous étudier un cas pratique où l'accès à  $h(t)$  se fait par l'intermédiaire de la réponse  $q(t)$  du système à une source  $v(t)$ . La relation de filtrage qui relie les quantités en jeu est :

$$q(t) = (h * v)(t). \quad (\text{A.28})$$

Le théorème des interférences amène à :

$$C_q(t) = (C_h * C_v)(t). \quad (\text{A.29})$$

Ainsi, en observant que :

$$\frac{d}{dt}C_q(t) = \frac{d}{dt}(C_h * C_v)(t) = (C_v * \frac{d}{dt}C_h)(t), \quad (\text{A.30})$$

nous en déduisons l'application pratique de l'identité de Ward, à savoir la relation entre la réponse impulsionnelle du système et la corrélation de la réponse de ce système, à une source donnée  $v(t)$  :

$$-4\eta L \dot{C}_q(t)u(t) = (C_v * h)(t). \quad (\text{A.31})$$



## Annexes du chapitre 2

---

### B.1 A propos de la dissipation visqueuse

L'air est un fluide visqueux, il résiste à sa mise en mouvement et donc à la propagation de la pression acoustique. L'opérateur de dissipation de type frottement fluide, illustré dans le cas de l'oscillateur amorti et amenant à l'identité (A.26), n'est pas adapté pour décrire le comportement de l'air. Nous rappelons dans un premier temps les phénomènes qui peuvent être modélisés pour constituer un opérateur de dissipation adapté.

Considérons la propagation d'un champ scalaire dans un milieu de propagation. Les hypothèses faites sur le milieu sont celles de **l'homogénéité** et de **l'isotropie**, pour les longueurs d'ondes mesurées ; c'est-à-dire que les propriétés du milieu (température, pression, composition chimique) sont les mêmes en tout point de l'espace et dans toutes les directions de propagation. Le milieu est également **linéaire**, définissant qu'une excitation monochromatique résulte en un champ identiquement monochromatique. L'opérateur de propagation est par ailleurs une forme linéaire et il préserve donc les combinaisons linéaires.

Premièrement, il est établi que les bords du milieu de propagation sont responsables d'une force de friction (contrainte de cisaillement) opposée à la direction de propagation du champ de pression. La norme de cette force dépend de la viscosité dynamique du fluide, qui vaut  $\mu = T_0^{0.73}$  dans l'air. L'application du principe fondamental de la dynamique permet d'établir l'équation de propagation, dans laquelle l'opérateur de dissipation prend la forme  $\alpha^2/c_0 \frac{\partial}{\partial t} \Delta$ , où  $c_0$  est la vitesse de propagation de l'onde.

Pour mesurer la pertinence de ce premier niveau de modélisation, considérons la propagation d'une onde plane harmonique se propageant en ligne droite. Appelons  $p(t, \mathbf{x})$  l'amplitude du champ scalaire en  $\mathbf{x}$  à l'instant  $t$ , réponse du milieu à la source excitatrice  $f(t, \mathbf{x})$ . L'équation de propagation d'un champ dans un milieu linéaire visqueux est proposée par Stokes [1],

$$\left[ \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} - \frac{\alpha^2}{c_0} \frac{\partial}{\partial t} \Delta - \Delta \right] p(t, \mathbf{x}) = f(t, \mathbf{x}). \quad (\text{B.1})$$

Dans le domaine des fréquences  $(\omega, k)$ , la relation de dispersion s'écrit

$$-k^2[1 + jk_0 l_v] + k_0^2 = 0, \text{ où } k_0 = \omega/c_0 \quad (\text{B.2})$$

Le vecteur d'onde est complexe,  $k = k' - jk''$  et  $k'$  est lié à la propagation tandis que  $k''$  est liée à l'atténuation. Etant donné que dans l'air et dans des conditions normales de température et de pression,  $k_0\alpha^2 \approx 2\omega \cdot 10^{-10}$  s est très faible jusqu'aux centaines de Hertz et une approximation au premier ordre du nombre d'onde est pertinente,

$$k = \frac{k_0}{\sqrt{1 + jk_0\alpha^2}} = k_0 - j\frac{\alpha^2}{2}k_0^2 + o(k_0^2\alpha^2) \quad (\text{B.3})$$

Un aspect intéressant vis-à-vis de la propagation, à cette étape de la modélisation, est que la vitesse de phase  $v_\phi = \omega/k' = c_0$  n'est pas perturbée par la viscosité. D'autre part, l'atténuation par unité de longueur s'exprime

$$k'' = \frac{l_v\omega^2}{2c_0^2}. \quad (\text{B.4})$$

Ainsi, l'atténuation augmente avec le carré de la fréquence, de telle sorte qu'à 1 kHz, l'amplitude de l'onde est atténuée de 1.3% tous les km, ce qui rend donc l'amortissement de cisaillement négligeable dans le domaine des fréquences audibles. Cependant pour les plus hautes fréquences, de l'ordre du MHz par exemple, la propagation est rendue très courte.

Le deuxième degré de modélisation du caractère visqueux de l'air consiste à prendre en compte sa nature de gaz, qui implique des pertes par conduction de la chaleur entre particules voisines et perte par relaxation thermique moléculaire. Pour rester synthétique, prendre en compte la conduction thermique revient à adapter la longueur caractéristique visqueuse  $\alpha^2$  pour la remplacer par la longueur caractéristique visco-thermique  $\alpha' \geq \alpha$ .

Pour un gaz polyatomique, comme c'est le cas de l'air, c'est la prise en compte de la relaxation thermique moléculaire qui achève de mettre à niveau le modèle théorique pour être en adéquation avec les absorptions mesurées en pratique. Lorsqu'un gaz subit une compression, la température s'élève, ce qui revient à dire que le mouvement de translations des molécules s'accélère. Cependant, les molécules diatomiques sont également animées par des mouvements de rotations et de vibrations. La translation excitatrice se transfère alors sous la forme de trois mouvements : translation, rotation et vibration. Il s'agit d'un transfert d'énergie, qui dans les deux derniers cas : translation/rotation et translation/vibration, ne contribue pas à la propagation de l'onde.

Le transfert translation/rotation est quasi-instantanée, et c'est le paramètre  $\eta$  qui en tient compte, ce qui revient à nouveau à augmenter la longueur caractéristique visco-thermique  $l_{vh}$ . En revanche, le transfert translation/vibration s'effectue avec un retard caractérisé par des temps de relaxation  $\theta_d$  dépendant des espèces chimiques  $d$  en jeu.

Dans l'air, dont la composition est approchée à 20% de  $O_2$  et 80% de  $N_2$ , alors

$$v_\phi = \frac{\omega}{k'} \approx c_0 \left( 1 + \frac{1}{2} \sum_q D_q \frac{(\omega\theta_q)^2}{1 + (\omega\theta_q)^2} \right) \approx c_0. \quad (\text{B.5})$$

L'atténuation par unité de longueur devient

$$k'' = \frac{\omega}{2c_0} \left( \frac{\omega l_{vh}}{c_0} + \sum_q D_q \frac{\omega \theta_q}{1 + (\omega \theta_q)^2} \right). \quad (\text{B.6})$$

L'atténuation par unité de longueur en fonction de la fréquence est présentée Fig. B.1. Elle se démarque de la simple évolution proportionnelle au carré de la fréquence par des sauts, correspondants à l'influence des espèces chimiques. Nous identifions cependant que dans la gamme de fréquences considérées [1 – 10] kHz, la prise en compte de la relaxation moléculaire est nécessaire pour rendre compte d'une modélisation réaliste, mais que le comportement est bien représenté par la loi d'évolution proportionnelle au carré de la fréquence. Nous retenons donc pour la suite que pour les fréquences considérées, la dissipation est modélisée par l'opérateur  $\alpha^2 \partial / \partial t \Delta$ , où  $\alpha^2$  tient compte de la dissipation thermo-visqueuse et de la relaxation moléculaire de  $N_2$  et  $O_2$ .

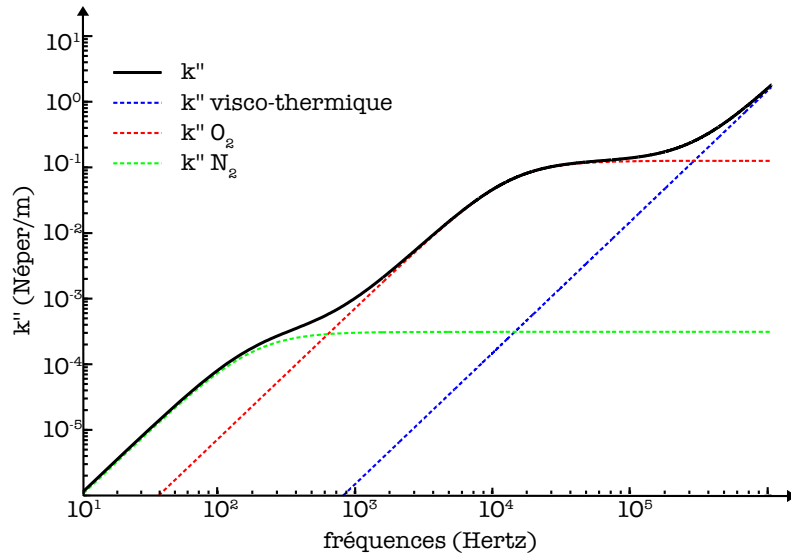


FIGURE B.1 – Atténuation par unité de longueur dans l'air en fonction de la fréquence.

En pratique, la propagation acoustique dans une salle est par ailleurs très fortement influencée par la géométrie de la salle et son encombrement (mobilier), générant à chaque réflexion spéculaire des couplages acoustico-mécaniques entraînant des pertes d'énergie conséquentes. Un ordre de grandeur à retenir pour ces salles, est la diminution de 60 dB de l'amplitude du champ en généralement moins d'une seconde. Dans une salle de dimension caractéristique 10 m, une trentaine de réflexions spéculaires sur le contour du milieu sont alors réalisées. En guise de comparaison, une excitation mécanique dans un diapason conduit à l'établissement de résonances qui perdurent plusieurs secondes, avec une onde qui se propage à plusieurs km/s et un couplage air/métal incessant. A l'inverse, dans du béton, le caractère fortement visqueux du matériau conduit à une vitesse de propagation de l'onde à 240 m/s, inférieure donc à la celle du son dans l'air, ainsi qu'à une atténuation de 60dB au bout de quelques mètres au plus.

## B.2 Preuve de l'identité de Ward

Plusieurs approches existent dans la littérature pour démontrer la relation (2.9). Celle que nous proposons ici exploite le caractère infini du milieu et positionne les calculs dans l'espace des fréquences  $(\omega, \mathbf{k})$ , simplifiant ainsi les relations de filtrage par de simples produits.

Pour commencer, partons de la définition de la corrélation de Green  $C_g$ ,

$$C_g = g * g_-^* \xleftrightarrow{TF(\omega, \mathbf{k})} \hat{C}_g(\omega, \mathbf{k}) = \hat{G}(\omega, \mathbf{k}) \hat{G}^*(\omega, \mathbf{k}) \quad (\text{B.7})$$

Or, dans le domaine des fréquences, l'équation de propagation s'écrit

$$(v^2 \mathbf{k}^2 + j\alpha^2 \mathbf{k}^2 \omega - \omega^2) \hat{P}(\omega, \mathbf{k}) = \hat{F}(\omega, \mathbf{k}) \quad (\text{B.8})$$

et la fonction de Green est par définition la réponse à une source impulsionnelle, ce qui dans l'espace des fréquences s'écrit  $\hat{F}(\omega, \mathbf{k}) = 1$ , pour une source d'énergie unitaire, soit donc

$$\hat{G}(\omega, \mathbf{k}) = \frac{1}{v^2 \mathbf{k}^2 - j\alpha^2 \omega \mathbf{k}^2 - \omega^2}. \quad (\text{B.9})$$

Pour relier corrélation de Green et fonction de Green, observons

$$(v^2 \mathbf{k}^2 - j\alpha^2 \omega \mathbf{k}^2 - \omega^2) \hat{C}_g(\omega, \mathbf{k}) = \hat{G}^*(\omega, \mathbf{k}). \quad (\text{B.10})$$

Ainsi, en égalisant les parties imaginaires des membres de droite et de gauche de l'égalité, on obtient l'identité de Ward dans le domaine des fréquences  $(\omega, \mathbf{k})$ ,

$$-j\alpha^2 \omega \mathbf{k}^2 \hat{C}_g(\omega, \mathbf{k}) = \frac{1}{2} \left( \hat{G}(\omega, \mathbf{k}) - \hat{G}^*(\omega, \mathbf{k}) \right). \quad (\text{B.11})$$

Nous souhaitons maintenant obtenir l'identité de Ward dans le domaine direct  $t, \mathbf{x}$ . Pour ce faire, prenons la TF inverse en  $\mathbf{k}$  de l'Eq (B.11). Cependant auparavant, exploitons l'approximation de faible viscosité de l'air afin de simplifier les calculs à venir, et posons  $\mathbf{k} \approx \omega/v$ . Alors, l'Eq. (B.11) devient

$$-j\alpha^2 \omega^3 \tilde{C}_g(\omega, \Delta \mathbf{x}) \approx \frac{1}{2} \left( \tilde{G}(\omega, \Delta \mathbf{x}) - \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \tilde{G}^*(\omega, \mathbf{k}) \exp((\cdot) - j\langle \mathbf{k}, \mathbf{x} \rangle) d\mathbf{k} \right) \quad (\text{B.12})$$

ce qui s'écrit également -

$$j\alpha^2 \omega^3 \tilde{C}_g(\omega, \Delta \mathbf{x}) \approx \frac{1}{2} \left( \tilde{G}(\omega, \Delta \mathbf{x}) - \left[ \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \tilde{G}(\omega, \mathbf{k}) \exp((\cdot) - j\langle \mathbf{k}, -\mathbf{x} \rangle) d\mathbf{k} \right]^* \right). \quad (\text{B.13})$$

Ainsi, dans le domaine  $(\omega, \mathbf{x})$ , l'identité de Ward s'écrit

$$-j\alpha^2 \omega \tilde{C}_g(\omega, \Delta \mathbf{x}) \approx \frac{1}{2} \left( \tilde{G}(\omega, \Delta \mathbf{x}) - \tilde{G}^*(\omega, -\Delta \mathbf{x}) \right) = \frac{1}{2} \left( \tilde{G}(\omega, \Delta \mathbf{x}) - \tilde{G}^*(\omega, \Delta \mathbf{x}) \right), \quad (\text{B.14})$$

car la fonction de Green en milieu infini est radiale et ne dépend que de  $\|\Delta \mathbf{x}\|$ , donc

$$-j\alpha^2\omega^3\tilde{C}_g(\omega, \Delta \mathbf{x}) \approx \mathbb{Im} \left( \tilde{G}(\omega, \Delta \mathbf{x}) \right). \quad (\text{B.15})$$

Finalement, en prenant la TF inverse en  $\omega$  et en identifiant que  $j\omega$  correspond à une dérivation, l'identité de Ward visco-acoustique dans le domaine  $(t, \mathbf{x})$  s'écrit

$$\alpha^2 \frac{\partial^3}{\partial t^3} C_g(\Delta t, \Delta \mathbf{x}) \approx \frac{1}{2} (g(\Delta t, \Delta \mathbf{x}) - g^*(-\Delta t, \mathbf{x})) = \frac{1}{2} (g(\Delta t, \Delta \mathbf{x}) - g(-\Delta t, \mathbf{x})) \quad (\text{B.16})$$

car la fonction de Green est réelle, et l'identité annoncée est bien retrouvée

$$\alpha^2 \frac{\partial^3}{\partial t^3} C_g(\Delta t, \Delta \mathbf{x}) \approx \text{Odd}(g(\Delta t, \Delta \mathbf{x})). \quad (\text{B.17})$$

□

### B.3 Propagation d'erreur par dérivation

La dérivation est un filtrage qui peut numériquement être implémenté avec des soustractions. Si l'on se donne un signal  $\mathbf{s}_m$ , version numérisée d'un champ  $s$  au temps  $mT_s$ , avec  $T_s$  la période d'échantillonnage, alors il est possible d'écrire :

$$\mathbf{s}_m = s(mT_s) + n_m \text{ où } n \sim \mathcal{N}(0, \sigma^2) \quad (\text{B.18})$$

La dérivée  $\alpha$ ième de  $\mathbf{s}_m$  s'écrit :

$$(\mathbf{s}_m)^\alpha = (\mathbf{s}_m)^{\alpha-1} - (\mathbf{s}_{m-1})^{\alpha-1} = s^\alpha(mT_s) + n_\alpha(0, \sigma_\alpha^2) \quad (\text{B.19})$$

Une conséquence directe est alors que le bruit  $n_\alpha$ , présent dans la version dérivée du signal, a des propriétés qui sont différentes de celles du bruit de mesure original  $n$ . Ces propriétés sont d'ailleurs une fonction de  $\alpha$ . Nous nous intéresserons maintenant à caractériser cette loi d'évolution de la variance  $\sigma_n^2$ .

Par (B.19),  $n_\alpha$  apparaît s'exprimer comme une suite auto-régressive de la séquence aléatoire i.i.d  $n$ . La variance de ce bruit est alors définie comme étant la somme des carrés des coefficients du modèle auto-régressif. Pour comprendre comment cette variance évolue avec  $\alpha$ , nous allons d'abord montrer par récurrence que

$$(\mathbf{s}_m)^\alpha = \sum_{k=0}^{\alpha} (-1)^k \binom{\alpha}{k} \mathbf{s}_{m-k} \quad (\text{B.20})$$

– Pour  $\alpha = 1$ , on remarque que :

$$(\mathbf{s}_m)^1 = (-1)^0 \binom{1}{0} \mathbf{s}_{m-0} + (-1)^1 \binom{1}{1} \mathbf{s}_{m-1} \quad (\text{B.21})$$



- Ensuite, supposons la propriété vraie au rang  $\alpha$ . Montrons qu'elle est alors toujours vraie au rang  $\alpha + 1$ . En fait, le développement de  $(\mathbf{s}_m)^{\alpha+1}$  en utilisant (B.19) et (B.20) s'écrit :

$$(\mathbf{s}_m)^{\alpha+1} = \sum_{k=0}^{\alpha} (-1)^k \binom{\alpha}{k} \mathbf{s}_{m-k} - \sum_{k=0}^{\alpha} (-1)^k \binom{\alpha}{k} \mathbf{s}_{m-k-1}$$

ce qui, en posant  $k' = k + 1$ , se simplifie en :

$$(\mathbf{s}_m)^{\alpha+1} = \sum_{k=0}^{\alpha+1} (-1)^k \left[ \binom{\alpha}{k} + \binom{\alpha}{k-1} \right] \mathbf{s}_{m-k} \quad (\text{B.22})$$

Il s'agit ensuite d'utiliser le fait que pour tout  $k$  :

$$\binom{\alpha}{k} + \binom{\alpha}{k-1} = \frac{\alpha!}{k!(\alpha-k)!} \times \frac{(\alpha-k+1)}{(\alpha-k+1)} + \frac{\alpha!}{(k-1)!(\alpha-k+1)!} \times \frac{k}{k} = \binom{\alpha+1}{k}$$

Ce qui achève la preuve par récurrence.

La contribution stochastique dans (B.20) est formée par la combinaison linéaire des bruits de mesure aux différents temps. Celui-ci est supposé être une séquence i.i.d gaussienne. En utilisant l'identité de Vandermonde il vient alors que :

$$\sigma_{\alpha}^2 = \left[ \sum_{k=0}^{\alpha} \binom{\alpha}{k}^2 \right] \times \sigma^2 = \binom{2\alpha}{\alpha} \times \sigma^2 \quad (\text{B.23})$$

□

## B.4 Probabilité d'erreur de détection de $t_1$

Calculons  $\mathbb{P}[w \geq A]$ , la probabilité qu'une variable aléatoire  $w$ , de densité  $p_w$  gaussienne centrée de variance  $\sigma^2$  soit plus grande que l'amplitude  $A$ . La relation de Chasles amène à :

$$\underbrace{\int_{-\infty}^{+\infty} p_w(t) dt}_{=1} = \underbrace{\int_{-\infty}^0 p_w(t) dt}_{=1/2} + \underbrace{\int_0^A p_w(t) dt}_{=I} + \underbrace{\int_A^{+\infty} p_w(t) dt}_{=\mathbb{P}[w \geq A]} \quad (\text{B.24})$$

où  $I = \frac{1}{2} \text{erf}\left(\frac{\sqrt{2}}{2} \frac{A}{\sigma}\right)$ . En effet,

$$\int_0^A p_w(t) dt = \frac{1}{\sqrt{2\pi}\sigma} \int_0^1 \exp\left(-\frac{t^2}{2\sigma^2}\right) dt \stackrel{t'=t/\sqrt{2}\sigma^2}{=} \frac{\sigma\sqrt{2}}{\sqrt{2\pi}\sigma} \underbrace{\int_0^{\frac{A}{\sigma\sqrt{2}}} \exp(-t'^2) dt'}_{=\frac{\sqrt{\pi}}{2} \text{erf}\left(\frac{A}{\sigma\sqrt{2}}\right)} \quad (\text{B.25})$$

ce qui donne bien le résultat présenté.

□

## B.5 Fonction de Green : $(\omega, \mathbf{x}) \rightarrow (t, \mathbf{x})$

Soit l'intégrale à calculer, par l'application du théorème des résidus :

$$A = TF_{\omega}^{-1}\{g(\omega, \mathbf{x}, \mathbf{y})\}[t, \mathbf{x}, \mathbf{y}] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} g(\omega, \mathbf{x}, \mathbf{y}) \exp(j\omega t) d\omega$$

Soit un contour  $\Gamma$  délimité par l'axe des abscisses  $\omega \in [-R, R]$  et par un arc de cercle  $\mathcal{C}$  décrit par  $z(k) = R \exp(jk)$ ,  $k \in [-\pi, \pi]$ . Ce contour est choisi tel qu'il entoure les pulsations propres  $\omega_n^{\pm}$ , lorsque  $R$  est suffisamment grand.

On calcule l'intégrale suivante, pour chaque composante  $n$  :

$$I_n = A_n + j \int_{\mathcal{C}} \langle g(z(k), \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle \exp(-jz(k)t) z(k) dk \quad (\text{B.26})$$

qui vaut aussi :

$$I_n = 2j\pi \sum_k \text{Res}(\langle g(z(k), \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle \exp(jzt), z_k). \quad (\text{B.27})$$

Le raisonnement ne sera mené que sur la  $n$ ème composante de la fonction de Green.

On montre, en utilisant le lemme de Jordan, que l'intégrale sur  $\mathcal{C}$  est nulle. Pour cela on pose  $f(z) = \langle g(z(k), \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle \exp(jz(k)t)$  et on évalue le comportement de  $|zf(z)|$  lorsque  $z$  se déplace sur le contour  $\mathcal{C}$  et quand  $R$  augmente vers l'infini :

$$|z(k)f(z(k))| = \frac{R}{|R \exp(jk - \omega_n^+)||R \exp(jk - \omega_n^-)|} \underset{R \rightarrow +\infty}{\sim} \frac{1}{R} \rightarrow 0 \quad (\text{B.28})$$

Il reste à calculer les résidus aux points singuliers de la fonction :

$$\begin{cases} \text{Res}(\langle g(z(k), \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle \exp(jzt), \omega_n^+) &= \lim_{z \rightarrow \omega_n^+} \frac{\exp(jkt)}{z - \omega_n^-} = + \frac{\exp(j\omega_n^+ t)}{q_n} \\ \text{Res}(\langle g(z(k), \mathbf{x}, \mathbf{y}), \mathbf{e}_n \rangle \exp(jzt), \omega_n^-) &= \lim_{z \rightarrow \omega_n^-} \frac{\exp(jkt)}{z - \omega_n^+} = - \frac{\exp(j\omega_n^- t)}{\omega_{\alpha, n}} \end{cases} \quad (\text{B.29})$$

Enfin, l'ensemble de ces résultats conduit à l'expression de  $I_n$  :

$$I_n 2\pi A = TF_{\omega}^{-1}\{g(\omega, \mathbf{x}, \mathbf{y})\}[t, \mathbf{x}, \mathbf{y}] = 2j\pi \frac{1}{\omega_{\alpha, n}} (\exp(j\omega_n^+ t) - \exp(j\omega_n^- t)) \quad (\text{B.30})$$

soit finalement :

$$\langle g(t, \mathbf{x}, \mathbf{y}), \mathbf{e}_n(\mathbf{x}) \rangle = -H(t) \frac{2}{\omega_{\alpha, n}} \exp\left(-\frac{d_n}{2}t\right) \sin(\omega_{\alpha, n} t) \mathbf{e}_n(\mathbf{y}) \quad (\text{B.31})$$

où  $H(t)$  est l'échelon de Heaviside.

□

## B.6 Corrélation de Green dans une base

Dans la base  $(\mathbf{e}_n)_{n \in \mathbb{N}}$ , la fonction de Green s'exprime

$$g(t, \mathbf{x}, \mathbf{y}) = H(t) \sum_{n \geq 0} \exp\left(-\frac{d_n}{2}t\right) \frac{\sin(q_n t)}{q_n} \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y}), \quad (\text{B.32})$$

et par définition, la corrélation de Green vaut

$$cg(\tau, \mathbf{x}, \mathbf{y}) = \int_{\mathbb{R}} \int_{\mathbb{X}} g(t, \mathbf{x}, \mathbf{s}) g(t + \tau, \mathbf{s}, \mathbf{y}) dt d\mathbf{s}. \quad (\text{B.33})$$

En combinant les deux équations précédentes, la corrélation de Green s'exprime :

$$\begin{aligned} cg(\tau, \mathbf{x}, \mathbf{y}) = \sum_{m, n \geq 0} \int_{\mathbb{R}} H(t) H(t + \tau) \exp\left(-\frac{d_m}{2}(t + \tau) - \frac{d_n}{2}t\right) \frac{\sin(q_n t)}{q_n} \frac{\sin(q_m(t + \tau))}{q_m} dt \\ \times \int_{\mathbb{X}} \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{s}) \mathbf{e}_m(\mathbf{s}) \mathbf{e}_m(\mathbf{y}) d\mathbf{s} \end{aligned} \quad (\text{B.34})$$

Pour mener le calcul, nous nous appuyons sur le fait que :

$$\bullet \quad \int_{\mathbb{X}} \mathbf{e}_n(\mathbf{s}) \mathbf{e}_m(\mathbf{s}) d\mathbf{s} = \delta_{n,m} \quad (\text{B.35})$$

$$\bullet \quad \int_{\mathbb{R}} H(t) H(t + \tau) f(t) dt = \int_{\max(0, -\tau)}^{+\infty} f(t) dt \quad (\text{B.36})$$

$$\bullet \quad \sin(q_n t) \sin(q_n(t + \tau)) = \frac{1}{2} (\cos(q_n \tau) - \cos(q_n(2t + \tau))) \quad (\text{B.37})$$

Il vient alors que la double somme devient une somme simple, où les indices  $n$  et  $m$  sont égaux. En remplaçant le produit des sinus l'expression suggérée et en précisant les bornes d'intégrations, la corrélation de Green devient :

$$\sum_{n \geq 0} \frac{\exp\left(-\frac{d_n}{2}\tau\right)}{2q_n^2} \underbrace{\int_{\max(0, -\tau)}^{+\infty} \exp(-d_n t) (\cos(q_n \tau) - \cos(q_n(2t + \tau))) dt}_{=I} \times \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y}) \quad (\text{B.38})$$

Pour clarifier les calculs, développons  $I$  :

$$I = \cos(q_n \tau) \underbrace{\int_{\max(0, -\tau)}^{+\infty} \exp(-d_n t) dt}_{=I_1} - \underbrace{\int_{\max(0, -\tau)}^{+\infty} \exp(-d_n t) \cos(q_n(2t + \tau)) dt}_{=I_2}. \quad (\text{B.39})$$

où

$$I_1 = \int_{\max(0, -\tau)}^{+\infty} \exp(-d_n t) dt = \left[ -\frac{\exp(-d_n t)}{d_n} \right]_{\max(0, -\tau)}^{+\infty} = \frac{\exp(-d_n \max(0, -\tau))}{d_n} \quad (\text{B.40})$$

et  $I_2$  se calcule par intégration par partie

$$I_2 = \underbrace{\left[ \exp(-d_n t) \frac{\sin(q_n(2t + \tau))}{2q_n} \right]_{\max(0, -\tau)}^{+\infty}}_{= -\frac{1}{2q_n} \exp(-d_n \max(0, -\tau)) \sin(q_n(2 \max(0, -\tau) + \tau))} + \underbrace{\frac{d_n}{2q_n} \int_{\max(0, -\tau)}^{+\infty} \exp(-d_n t) \sin(q_n(2t + \tau)) dt}_{= I_3} \quad (\text{B.41})$$

et

$$I_3 = -\frac{d_n}{2q_n} \underbrace{\left[ \exp(-d_n t) \frac{\cos(q_n(2t + \tau))}{2q_n} \right]_{\max(0, -\tau)}^{+\infty}}_{= \frac{d_n}{4q_n^2} \exp(-d_n \max(0, -\tau)) \cos(q_n(2 \max(0, -\tau) + \tau))} - \frac{d_n^2}{4q_n^2} I_2. \quad (\text{B.42})$$

Sachant que  $\tau/2 + \max(0, -\tau) = |\tau|/2$ , il vient :

$$\exp\left(-\frac{d_n}{2}\tau\right) I_2 \left(1 + \frac{d_n^2}{4q_n^2}\right) = \exp\left(-\frac{d_n}{2}|\tau|\right) \left( \frac{d_n}{4q_n^2} \cos(q_n|\tau|) - \frac{1}{2q_n} \sin(q_n|\tau|) \right) \quad (\text{B.43})$$

soit encore

$$\exp\left(-\frac{d_n}{2}\tau\right) I_2 (4q_n^2 + d_n^2) = \exp\left(-\frac{d_n}{2}|\tau|\right) (d_n \cos(q_n|\tau|) - 2q_n \sin(q_n|\tau|)) \quad (\text{B.44})$$

Ainsi, la corrélation de Green s'exprime :

$$cg(\tau, \mathbf{x}, \mathbf{y}) = \sum_{n \geq 0} \frac{\exp\left(-\frac{d_n}{2}|\tau|\right)}{2q_n^2} \left( \left( \frac{1}{d_n} - \frac{d_n}{4q_n^2 + d_n^2} \right) \cos(q_n\tau) + \frac{2q_n}{4q_n^2 + d_n^2} \sin(q_n|\tau|) \right) \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y}) \quad (\text{B.45})$$

ce qui amène à l'expression finale

$$cg(\tau, \mathbf{x}, \mathbf{y}) = \sum_{n \geq 0} \frac{\exp\left(-\frac{d_n}{2}|\tau|\right)}{(4q_n^2 + d_n^2)d_n q_n} (2q_n \cos(q_n\tau) + d_n \sin(q_n|\tau|)) \mathbf{e}_n(\mathbf{x}) \mathbf{e}_n(\mathbf{y}) \quad (\text{B.46})$$

## B.7 Equivalence de sources

La relation entre le temps de propagation inter-capteur et la différence de marche dépend de la géométrie source-capteurs. Cette relation n'est pas une bijection, puisque plusieurs

sources peuvent conduire à la même différence de marche ; en effet, par définition, si l'on décrit l'ensemble  $\mathcal{H}$  tel que

$$\forall \mathbf{s} \in \mathcal{H}, \quad d(\mathbf{s}, \mathbf{x}) - d(\mathbf{s}, \mathbf{y}) = \Delta, \quad (\text{B.47})$$

alors  $\mathcal{H}$  est l'hyperbole dont les foyers sont  $\mathbf{x}$  et  $\mathbf{y}$ , d'équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (\text{B.48})$$

Le paramètre  $a$  se détermine en considérant l'intersection de l'hyperbole avec le segment  $[\mathbf{x}; \mathbf{y}]$ , de telle sorte que  $a = \Delta/2$ . Le paramètre  $b$  est alors calculable en injectant  $a$  dans l'équation de l'hyperbole, ce qui donne  $b = \frac{1}{2} \sqrt{(c_0 t_1 + \Delta)(c_0 t_1 - \Delta)}$ .

Finalement, en considérant l'asymptote de l'hyperbole, dont la pente est  $\pm b/a$ , nous pouvons définir l'angle d'incidence  $\theta_{eq}$  convoité

$$\theta_{eq} = \arctan \left( \frac{1}{2\Delta} \sqrt{(c_0 t_1 + \Delta)(c_0 t_1 - \Delta)} \right). \quad (\text{B.49})$$

□

## Annexes du chapitre 3

---

### C.1 Détecteur d'impulsion acoustique

L'étage analogique de détection d'impulsion acoustique a pour rôle de d'avertir le micro-contrôleur lorsqu'une impulsion est détectée. Au réveil du noeud, seul cet étage est actif et guète l'arrivée d'un signal utile. Lorsqu'un tel évènement s'est produit, l'étage est éteint et n'a plus aucun rôle jusqu'au cycle de réveil suivant. Le microphone passif fourni un signal centré et dont l'amplitude dépend de la puissance de la source. De manière générale, ce signal évolue dans une plage qui dépasse rarement la centaine de mV avec un microphone passif, en observant des évènements acoustiques d'opportunité de type « claquement de main ». Un gain destiné à utiliser la dynamique complète proposée par les AOP basse consommation (typiquement  $\pm 3.3V$ ) peut être déployé. Pour détecter une impulsion acoustique, une forte variation de l'énergie du signal est attendue. L'étage analogique est donc composé d'un redresseur-intégrateur pour obtenir une image de l'énergie instantanée du champ. La constante de temps de l'intégrateur est choisie en adéquation avec la dynamique des signaux impulsionnels acoustiques, soit 0.01 sec. Pour imposer la détection d'un évènement brut, un passe-haut (par exemple de fréquence de coupure  $f_c = 100$  Hz), assure de ne pas déclencher l'acquisition lorsque le niveau sonore augmente progressivement et ne correspond pas à un évènement impulsionnel. Un comparateur vient alors définir un seuil définissant la sensibilité de la détection de l'impulsion. Le signal en sortie peut attaquer le micro-contrôleur sous forme d'un signal TTL, pour l'avertir de l'arrivée d'un évènement acoustique impulsionnel.

Le schéma présenté Fig. C.2, construit avec des composants basse consommation, amène à une consommation stable de quelques  $\mu W$ . En guise de comparaison, en prenant l'exemple d'une pile classique qui possède une capacité de charge de 175mAh, l'étage analogique seul consomme typiquement 1  $\mu A$  en fonctionnement, ce qui a très peu d'influence sur l'autonomie du noeud. La Fig. C.1 présente un profil des signaux à divers endroits du montage, assurant son bon fonctionnement.

### C.2 Quantification 1 bit : dérivation d'un critère de fin de coda

Soient deux variables aléatoires  $A$  et  $B$  Gaussiennes centrées de variances respectives  $\sigma_A^2$  et  $\sigma_B^2$ .  $A$  représente la coda, le signal utile, tandis que  $B$  est le bruit de mesure. L'objectif de cette note est de définir une règle pour choisir le temps de fin du signal de coda. Le critère

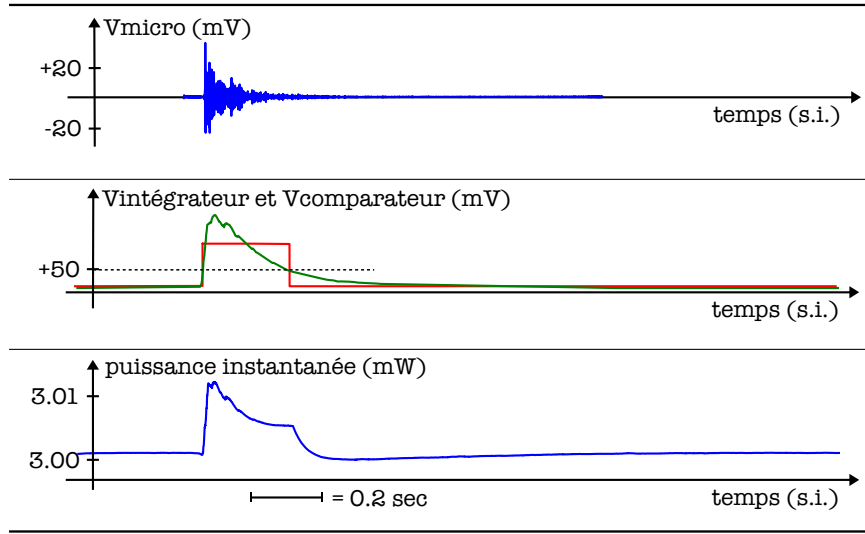


FIGURE C.1 – Résultats de l'étage analogique AID « Acoustic Impulse Detector »

proposé repose sur le fait que la mesure disponible est le signe de la somme des deux variables.

Nous avons :

$$\mathbb{P}[s(A+B) = s(A)] = \mathbb{P}[s(A) = s(B)] + \mathbb{P}[|A| > |B| \cap s(A) \neq s(B)] \quad (\text{C.1})$$

et étant donné que  $A$  et  $B$  suivent des lois symétriques, cette expression se réduit à :

$$\mathbb{P}[s(A+B) = s(A)] = \frac{1}{2} + \frac{1}{2}\mathbb{P}[|A| > |B|]. \quad (\text{C.2})$$

Il s'agit donc de calculer  $\mathbb{P}[|A| - |B| \geq 0]$  :

$$\begin{aligned} \mathbb{P}[|A| - |B| \geq 0] &= \int_0^{+\infty} p_{|A|-|B|}(x) dx \\ &= \int_0^{+\infty} \left[ \int_{\mathbb{R}} p_{|A|}(x-y) p_{|B|}(y) dy \right] dx \\ &\stackrel{y'=-y}{=} \int_0^{+\infty} \left[ \int_{\mathbb{R}} p_{|A|}(x+y') p_{|B|}(y') dy' \right] dx \\ \mathbb{P}[|A| - |B| \geq 0] &\stackrel{x'=x+y'}{=} \int_{\mathbb{R}} p_{|B|}(y') \left[ \int_{y'}^{+\infty} p_{|A|}(x') dx' \right] dy' \end{aligned} \quad (\text{C.3})$$

Dans cette expression, il s'agit maintenant de faire apparaître les densités des variables  $A$  et  $B$  et non plus celles de  $|A|$  et  $|B|$ . Une astuce de calcul apparaît en exploitant le fait que ces deux variables ont une densité de probabilité centrée symétrique, ce qui permet d'écrire :

$$\mathbb{P}[|A| - |B| \geq 0] = 4 \int_0^{+\infty} p_B(y') \int_{y'}^{+\infty} p_A(x') dx' dy' \quad (\text{C.4})$$

La combinaison des équations (C.2) et (C.4) pour des lois Gaussiennes centrées mène à :

$$\mathbb{P}[s(A+B) = s(A)] = \frac{1}{2} + \frac{2}{2\sigma_A\sigma_B\pi} \int_0^{+\infty} \exp\left(\frac{-y'^2}{2\sigma_B^2}\right) \int_{y'}^{+\infty} \exp\left(\frac{-x'^2}{2\sigma_A^2}\right) dx' dy', \quad (\text{C.5})$$

où il est reconnaissable que :

$$\int_{y'}^{+\infty} \exp\left(\frac{-x'^2}{2\sigma_A^2}\right) dx' = \sigma_A \sqrt{\frac{\pi}{2}} (1 - \operatorname{erf}[\frac{y'}{\sigma_A \sqrt{2}}]). \quad (\text{C.6})$$

Par ailleurs, en posant  $y = y'/\sigma_B \sqrt{2}$ , l'expression simplifiée suivant apparaît :

$$\mathbb{P}[s(A+B) = s(A)] = \frac{1}{2} + \mathbb{P}[B \geq 0] - \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \operatorname{erf}[\frac{\sigma_B}{\sigma_A} y] \exp(-y^2) dy. \quad (\text{C.7})$$

L'étape suivante est le calcul de la quantité :  $\Delta = \int_0^{t \geq 0} \operatorname{erf}[\alpha y] \exp(-y^2) dy$ . Pour ce faire, nous allons d'abord la dériver par rapport à  $\alpha$ , pour évaluer la limite lorsque  $t$  tend vers l'infini. L'intégration en  $\alpha$  permettra à la fin de remonter à la quantité souhaitée. La dérivation s'exprime à l'aide d'intégration par partie, de telle sorte que :

$$\frac{\partial}{\partial \alpha} \Delta = \frac{2}{\sqrt{\pi}} \int_0^{t \geq 0} y \exp(-(1+\alpha^2)y^2) dy = \frac{2}{\sqrt{\pi}} \frac{1 - \exp(-(1+\alpha^2)t^2)}{2(1+\alpha^2)}. \quad (\text{C.8})$$

La limite lorsque  $t$  augmente à l'infini vaut alors :

$$\frac{\partial}{\partial \alpha} \Delta \xrightarrow{t \rightarrow \infty} \frac{1}{\sqrt{\pi}(1+\alpha^2)}. \quad (\text{C.9})$$

Finalement, l'intégration par rapport à  $\alpha$  pour revenir à  $\Delta$  s'exprime :

$$\mathbb{P}[s(A+B) = s(A)] = 1 - \frac{1}{\sqrt{\pi}} \left( \frac{1}{\sqrt{\pi}} \arctan(\alpha) + k \right) \quad (\text{C.10})$$

où  $k$  est la constante d'intégration, choisie nulle, car pour  $\alpha = 0$  (un RSB infini), il est attendu du signe de  $A+B$  d'être tout le temps égal au signe de  $A$ .

### Résultat C.1

*La relation à retenir est donc que pour deux variables  $A$  et  $B$  gaussiennes centrées :*

$$\operatorname{signe}(A+B) = (-1)^\lambda \times \operatorname{signe}(A) \quad (\text{C.11})$$

où  $\lambda$  est une variable aléatoire de Bernoulli de paramètre  $p = \frac{1}{\pi} \arctan \frac{\sigma_B}{\sigma_A}$ .

A partir de ce résultat, l'utilisateur peut définir un seuil sur le RSB pour contrôler la probabilité que le signe de  $A+B$  soit différent du signe de  $A$ . Un critère signalant le temps à partir duquel la coda n'est plus prédominante par rapport au bruit de mesure en émerge.

### Exemple C.1

*Considérons le cas de la corrélation de signaux de codas quantifiées sur 1 bit. Cette corrélation s'exprime comme la somme du produit des signes. On peut donc par exemple définir un seuil  $\epsilon$  pour chacun des produits tel que  $\forall i, \mathbb{P}[s(A_i + B_i)s(A'_i + B'_i) - s(A_i)s(B_i) \leq \epsilon]$ , auquel cas il faudra vérifier que :*

$$\left( \frac{1}{\pi} \arctan \frac{\sigma_B}{\sigma_A} - 1 \right) \left( \frac{1}{\pi} \arctan \frac{\sigma_{B'}}{\sigma_{A'}} - 1 \right) \geq 1 - \epsilon \quad (\text{C.12})$$





FIGURE C.2 – Exemple de schéma d'un détecteur d'impulsion

# Annexes du chapitre 4

## D.1 Modèle polynomial par morceaux de la corrélation

L'identité de Ward relie la corrélation de Green à la partie impaire de la fonction de Green. Par linéarité, l'intégration de la fonction de Green qui est une somme de retards est également la somme de l'intégration des retards. L'intégration 3 fois d'un unique temps de propagation  $t_i$  associé à une amplitude  $a_i$ , est représenté dans la Fig. D.1. En considérant la parité des fonctions considérées, la corrélation de Green s'écrit :

$$cg(t \geq 0, \mathbf{x}, \mathbf{y} | t_i, a_i) = [\alpha t^2 + \beta t + \gamma](1 - H(t - t_i)) + [\epsilon t + \delta]H(t - t_i) \quad (D.1)$$

où

$$\begin{aligned} \alpha &= a_i \\ \beta &= a_i t_i \\ \gamma &= a_i t_i^2 + \text{offset}_i \\ \epsilon &= -3a_i t_i \\ \delta &= \text{offset}_i \end{aligned} \quad (D.2)$$

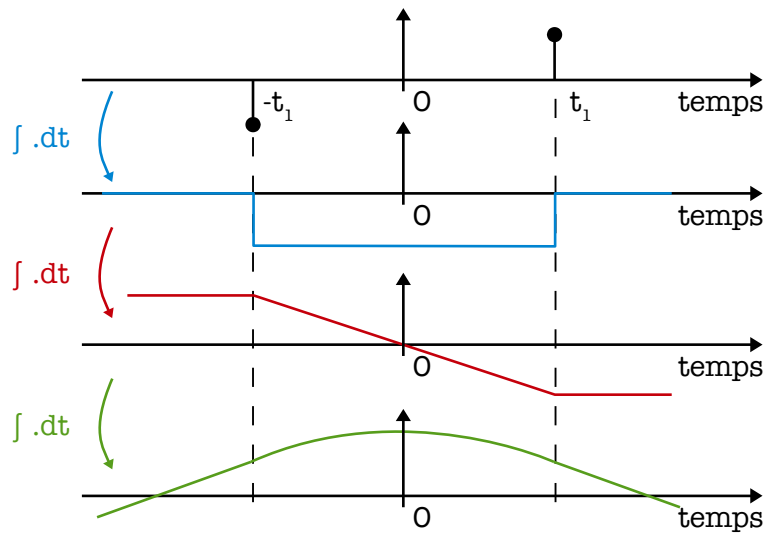


FIGURE D.1 – Modèle polynomial par morceaux de la corrélation de Green

En théorie, la fonction de Green est composée d'une infinité de temps de propagation. Par linéarité, le ième morceau est la superposition des morceaux associés aux temps de propagations qui sont plus grands :

$$cg(t, \mathbf{x}, \mathbf{y} | t_i \geq t \geq t_{i+1}) = \left( \sum_{k \geq i} \alpha_k \right) t^2 + \left( \sum_k \beta_k \right) t + \sum_k \gamma_k \quad (\text{D.3})$$

en conséquence de quoi il est notable que le premier morceau dépend de l'ensemble des temps de propagation. Quoiqu'il en soit, la corrélation de Green peut bien être modélisée par le modèle plus simple présenté Eq. (4.3).

□

## D.2 Exemple de marginalisation

Pour illustrer la notion de marginalisation, considérons la variable aléatoire  $y = v + w$ , où  $v$  et  $w$  sont deux lois normales centrées indépendantes, de variances  $\sigma_v^2$  et  $\sigma_w^2$  respectivement. La variable  $y$  a pour densité :

$$p_y(t') = \int_{\mathbb{R}} p_v(t) p_w(t' - t) dt = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{\sigma_v} \frac{1}{\sigma_w} \exp\left(-\frac{t^2}{2\sigma_v^2}\right) \exp\left(-\frac{(t' - t)^2}{2\sigma_w^2}\right) dt \quad (\text{D.4})$$

L'argument de l'exponentielle contient :

$$\frac{\sigma_w^2 t'^2 + \sigma_v^2 (t - t')^2}{\sigma_w^2 \sigma_v^2} = \frac{\sigma_w^2 + \sigma_v^2}{\sigma_w^2 \sigma_v^2} \left( t - \frac{\sigma_v^2 t'}{\sigma_w^2 + \sigma_v^2} \right)^2 + \frac{t'^2}{\sigma_w^2 + \sigma_v^2} \quad (\text{D.5})$$

Alors,

$$p_y(t') = \frac{\exp\left(-\frac{1}{2} \frac{t'^2}{\sigma_w^2 + \sigma_v^2}\right)}{\sqrt{2\pi(\sigma_v^2 + \sigma_w^2)}} \underbrace{\int_{\mathbb{R}} \sqrt{\frac{\sigma_v^2 + \sigma_w^2}{2\pi\sigma_v^2\sigma_w^2}} \exp\left(-\frac{1}{2} \frac{\sigma_w^2 + \sigma_v^2}{\sigma_w^2 \sigma_v^2} \left( t - \frac{\sigma_v^2 t'}{\sigma_w^2 + \sigma_v^2} \right)^2\right) dt}_{=1} \quad (\text{D.6})$$

ce qui permet de vérifier que  $y$  est une loi normale centrée de variance  $\sigma_v^2 + \sigma_w^2$ . Ce qu'il faut retenir, c'est qu'identifier une densité de probabilité connue et l'intégrer sur le domaine complet a permis (l'intégrale valant 1), de s'en défaire complètement.

## D.3 Marginalisation de la loi cible

La finalité est d'obtenir la distribution postérieure des temps de propagation, sachant le nombre de temps recherchés. C'est le facteur de Bayes qui sera déployé pour sélectionner le modèle (nombre de temps de propagation), le plus probable. La distribution cible est donc

$$p(\mathbf{t}, \boldsymbol{\beta}, \sigma^2, g | y) \propto \mathcal{L}(y | \mathbf{t}, \boldsymbol{\beta}, \sigma^2, g) \times p(t, \boldsymbol{\beta} | \sigma^2, g) \times p(\sigma^2, g). \quad (\text{D.7})$$

Etant données les relations de dépendance entre les paramètres, la distribution cible s'écrit également

$$p(\mathbf{t}, \boldsymbol{\beta}, \sigma^2, g|y) \propto \mathcal{L}(y|\mathbf{t}, \boldsymbol{\beta}, \sigma^2) \times p(t)p(\boldsymbol{\beta}|\sigma^2, g) \times p(\sigma^2)p(g). \quad (\text{D.8})$$

Etant donnée la finalité visée, le jeu de paramètres  $[\boldsymbol{\beta}, \sigma^2, g]$  est à marginaliser dans la distribution cible. Ici la marginalisation est en commençant par  $\boldsymbol{\beta}$ , puis  $\sigma^2$ . Comme on le verra, le g-prior ne pourra pas être marginalisé.

### D.3.1 Marginalisation de $\boldsymbol{\beta}$

La première marginalisation s'opère comme suit :

$$\begin{aligned} p(\mathbf{t}, \sigma^2, g|y) &= \int p(\mathbf{t}, \boldsymbol{\beta}, \sigma^2, g|y) d\boldsymbol{\beta} \\ &\propto p(t)p(\sigma^2)p(g) \int \mathcal{L}(y|\mathbf{t}, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}|\sigma^2, g) d\boldsymbol{\beta}. \end{aligned} \quad (\text{D.9})$$

où

$$\begin{aligned} -\mathcal{L}(y|\mathbf{t}, \boldsymbol{\beta}, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}^n} \times \exp\left(-\frac{1}{2\sigma^2}(S\boldsymbol{\beta} - \mathbf{y})^T \Lambda^{-1}(S\boldsymbol{\beta} - \mathbf{y})\right) \\ -p(\boldsymbol{\beta}|\sigma^2, g) &= \frac{1}{\sqrt{2\pi g\sigma^2}^{3K}} \times \exp\left(-\frac{1}{2g\sigma^2}(S\boldsymbol{\beta})^T \Lambda^{-1}(S\boldsymbol{\beta})\right) \end{aligned}$$

On observe alors le produit de deux densité Gaussiennes. Par identification, on cherche alors la moyenne équivalente  $\bar{\boldsymbol{\beta}}$  et la matrice de covariance équivalente  $\bar{A}$ , sachant que

$$(\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})^T A^{-1}(\boldsymbol{\beta} - \bar{\boldsymbol{\beta}}) = \boldsymbol{\beta}^T A^{-1}\boldsymbol{\beta} - 2\bar{\boldsymbol{\beta}}^T A^{-1}\boldsymbol{\beta} + \bar{\boldsymbol{\beta}}^T A^{-1}\bar{\boldsymbol{\beta}} \quad (\text{D.10})$$

est à rendre équivalent à

$$\mathbf{y}^T \Lambda^{-1} \mathbf{y} - 2\mathbf{y}^T \Lambda^{-1} S\boldsymbol{\beta} + \frac{1+g}{g}(S\boldsymbol{\beta})^T \Lambda^{-1}(S\boldsymbol{\beta}). \quad (\text{D.11})$$

On trouve alors que

$$\begin{aligned} -A &= \frac{g}{1+g} \times [S^T \Lambda^{-1} S]^{-1} \\ -\bar{\boldsymbol{\beta}} &= A S^T \Lambda^{-1} \mathbf{y} \end{aligned}$$

de telle sorte que

$$\mathcal{L}(y|\mathbf{t}, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta}|\sigma^2, g) \propto \lambda \times \left(\frac{1}{\sigma^2}\right)^{\frac{3K}{2}} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})^T A^{-1}(\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})\right) \quad (\text{D.12})$$

où

$$\lambda = \left(\frac{1}{\sigma^2}\right)^{\frac{N}{2}} \left(\frac{1}{g}\right)^{\frac{3K}{2}+1} \exp\left(-\frac{1}{2\sigma^2} \mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_N - \frac{g}{1+g} S[S^T \Lambda^{-1} S]^{-1} S^T \Lambda^{-1}\} \mathbf{y}\right). \quad (\text{D.13})$$

Ainsi, l'intégrale porte sur la densité d'une Gaussienne non normalisée, de telle sorte que le résultat de la marginalisation vaut

$$p(\mathbf{t}, \sigma^2, g|y) \propto \lambda \times \left(\frac{g}{1+g}\right)^{\frac{3K}{2}} \quad (\text{D.14})$$

### D.3.2 Marginalisation de $\sigma^2$

Pour marginaliser suivant  $\sigma^2$ , on observe que

$$\begin{aligned} p(\mathbf{t}, g | \mathbf{y}) &= \int p(\mathbf{t}, \sigma^2, g | \mathbf{y}) d\sigma^2 \\ &\propto p(\mathbf{t}) p(g) \int \lambda \times p(\sigma^2) \times \left( \frac{g}{1+g} \right)^{\frac{3K}{2}} d\sigma^2. \end{aligned} \quad (\text{D.15})$$

où  $p(\sigma^2) \sim 1/\sigma^2 \mathbb{1}(\mathbb{R}^+)$ , forme un a priori non informatif. Il est alors possible d'identifier la densité d'une loi inverse gamma pour la variable  $\sigma^2$ , de paramètres

$$\begin{aligned} - a &= N/2 \\ - b &= \frac{1}{2}(\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_n - \frac{g}{1+g} S[S^T \Lambda^{-1} S]^{-1} S^T \Lambda^{-1}\} \mathbf{y}) \end{aligned}$$

à laquelle il manque le facteur de normalisation  $b^a/\Gamma(a)$ . Son intégrale valant l'unité, après calculs, il reste finalement

$$p(\mathbf{t}, g | \mathbf{y}) \propto \frac{p(g) p(\mathbf{t}) g^{-1} (1+g)^{-\frac{3K}{2}}}{\left[ \mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_n - \frac{g}{1+g} S[S^T \Lambda^{-1} S]^{-1} S^T \Lambda^{-1}\} \mathbf{y} \right]^{\frac{N}{2}}}. \quad (\text{D.16})$$

## D.4 Mise à jour des paramètres : collapsed Gibbs sampling

A chaque mouvement, de nouveaux échantillons sont tirés pour estimer la distribution des paramètres. Une fois les temps de rupture mis à jour, le modèle polynomial par morceaux est à mettre à jour. Plus précisément, les paramètres du modèle, dont les lois dépendent de la position des ruptures, doivent évoluer. Il est possible de décrire les lois postérieures des paramètres  $\sigma^2$ ,  $\beta$  et  $g$  à un paramètre multiplicatif près ; pour cette raison, c'est un algorithme d'échantillonnage de Gibbs qui sera utilisé pour tirer des échantillons des lois obtenues.

### D.4.1 Mise à jour de $\sigma^2$

La loi postérieure de  $\sigma^2$  s'écrit

$$\begin{aligned} p(\sigma^2 | \mathbf{y}, \mathbf{t}, g) &= p(\mathbf{y}, \mathbf{t}, g | \sigma^2) p(\sigma^2) / p(\mathbf{y}, \mathbf{t}, g) \\ &= p(\mathbf{y} | \sigma^2, \mathbf{t}, g) p(\mathbf{t}, g | \sigma^2) p(\sigma^2) / p(\mathbf{y}, \mathbf{t}, g) \\ &\propto p(\mathbf{y} | \sigma^2, \mathbf{t}, g) p(\sigma^2) \end{aligned} \quad (\text{D.17})$$

auquel cas on identifie alors que  $\sigma^2 | \mathbf{y}, \mathbf{t}, g$  suit une loi inverse Gamma

$$\sigma^2 | \mathbf{y}, \mathbf{t}, g \sim \mathcal{IG}\left(\frac{n}{2}, \frac{1}{2}(\mathbf{y}^T \Lambda^{-1} \{\mathbb{1}_n - \frac{g}{1+g} S[S^T \Lambda^{-1} S]^{-1} S^T\} \mathbf{y})\right) \quad (\text{D.18})$$

D.4.2 Mise à jour de  $\beta$ 

La loi postérieure de  $\beta$  s'écrit

$$\begin{aligned}
 p(\beta|\mathbf{y}, \sigma^2, \mathbf{t}, g) &= p(\mathbf{y}, \mathbf{t}, g, \sigma^2|\beta)p(\beta)/p(\mathbf{y}, \mathbf{t}, g, \sigma^2) \\
 &= p(\mathbf{y}|\beta, \mathbf{t}, g, \sigma^2)p(\mathbf{t}, g, \sigma^2|\beta)p(\beta|\sigma^2, g, \mathbf{t})/p(\mathbf{y}, \mathbf{t}, g, \sigma^2) \\
 &= p(\mathbf{y}|\beta, \mathbf{t}, g, \sigma^2)p(\mathbf{t}, g, \sigma^2)p(\beta|\sigma^2, g, \mathbf{t})/p(\mathbf{y}, \mathbf{t}, g, \sigma^2) \\
 &\propto p(\mathbf{y}|\beta, \mathbf{t}, g, \sigma^2)p(\beta|\sigma^2, g, \mathbf{t})
 \end{aligned} \tag{D.19}$$

auquel cas on identifie alors que  $\sigma^2|\mathbf{y}, \mathbf{t}, g, \beta$  suit une loi Gaussienne

$$\beta|\mathbf{y}, \sigma^2, \mathbf{t}, g \sim \mathcal{N}\left(\frac{g}{1+g}S[S^T\Lambda^{-1}S]^{-1}S^T\Lambda^{-1}\mathbf{y}, \sigma^2\frac{g}{1+g}S[S^T\Lambda^{-1}S]^{-1}\right) \tag{D.20}$$

D.4.3 Mise à jour de  $g$ 

La loi postérieure de  $g$  s'écrit

$$\begin{aligned}
 p(g|\sigma^2, \mathbf{t}, \beta) &= p(\mathbf{t}, \sigma^2, \beta|g)p(g)/p(\mathbf{t}, \beta, \sigma^2) \\
 &= p(\beta|g, \mathbf{t}, \sigma^2)p(\mathbf{t}, \sigma^2|g)p(g)/p(\mathbf{t}, \beta, \sigma^2) \\
 &\propto p(\beta|g, \mathbf{t}, \sigma^2)p(g)
 \end{aligned} \tag{D.21}$$

auquel cas on identifie alors que  $g|\sigma^2, \mathbf{t}, \beta$  suit une loi inverse Gamma

$$g|\sigma^2, \mathbf{t}, \beta \sim \mathcal{IG}\left(A + \frac{3K}{2}, B + \frac{1}{2\sigma^2}\beta^T S^T \Lambda^{-1} S \beta\right) \tag{D.22}$$

## D.5 Corrélation dans une marche aléatoire

Soit  $\mathbf{x}$  une séquence correspondant à une marche aléatoire gaussienne :

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathcal{N}(0, \sigma^2) = \mathbf{x}_0 + \mathcal{N}(0, i\sigma^2) \tag{D.23}$$

Pour évaluer la corrélation entre des échantillons de la chaîne, on peut utiliser le coefficient de Pearson :

$$\rho(\mathbf{x}_{i+k}, \mathbf{x}_i) = \frac{\text{cov}(\mathbf{x}_{i+k}, \mathbf{x}_i)}{\text{std}(\mathbf{x}_{i+k})\text{std}(\mathbf{x}_i)} = \sqrt{\frac{i}{i+k}} \tag{D.24}$$

où  $\text{cov}(\mathbf{x}_{i+k}, \mathbf{x}_i) = \mathbb{E}[(\mathbf{x}_{i+k} - \mathbf{x}_0)(\mathbf{x}_i - \mathbf{x}_0)] = i\sigma^2$ .

**Résultat D.1**

*Ainsi, si l'on souhaite que chacun des échantillons qui constituent la densité cible estimée soit « décorrélié » du précédent, un seuil  $\epsilon$  sur le coefficient de Pearson indique que l'écart  $k$  entre deux échantillons acceptés doit vérifier :*

$$k \geq i\left(\frac{1}{\epsilon^2} - 1\right) \tag{D.25}$$

Par exemple, si l'on souhaite imposer un coefficient de Pearson constant inférieur à 0.5, l'écart entre deux échantillons acceptés va croître linéairement jusqu'à 3000 lorsque le 1000ème échantillon sera sélectionné. Il aura fallu 4501500 mouvements de chaîne acceptés.

A l'inverse, si  $k$  est laissé constant égal à 100, le coefficient de Pearson atteint 0.5 au 35ème échantillon, est supérieur à 0.9 au 427 et vaut 0.95 au 1000ème. Il aura fallu 100000 mouvements de chaîne acceptés.

# A propos du décrément aléatoire

---

## E.1 Quantités de Green et décrément aléatoire

Rappelons l'expression du décrément aléatoire pour une variable  $X$  :

$$D_{XX}(t) = \mathbb{E} [X(t_0 + t)|X(t_0) = a, X'(t_0) = b] = a \frac{C_{XX}(t)}{\sigma_X^2} - b \frac{C'_{XX}(t)}{\sigma_{X'}^2}. \quad (\text{E.1})$$

Chacun des deux termes de droite informe sur une quantité différente, cependant liée au système. Chacune peut être isolé ; en effet :

- Si  $a = cst \neq 0$  et  $b$  non contraint (variable aléatoire centrée), alors :

$$D_{XX}(t) \xrightarrow{n \rightarrow \infty} a \frac{C_{XX}(t)}{\sigma_X^2} - \mathbb{E}[b] \frac{C'_{XX}(t)}{\sigma_{X'}^2} \propto C_{XX}(t) \quad (\text{E.2})$$

auquel cas les paramètres caractéristiques du système sont identifiables car liés à ceux du décrément aléatoire (même pulsation, carré de l'amortissement). C'est l'approche classiquement usitée. Un estimateur possible est alors :

$$\hat{D}_{XX}^1(t) = \frac{1}{K} \sum_{k \geq 1} (X(t_k + t)|X(t_k) = a \neq 0); \quad (\text{E.3})$$

- Si  $a = 0$  et  $b$  toujours de même signe, par exemple  $b \geq 0$ , alors :

$$D_{XX}(t) \xrightarrow{n \rightarrow \infty} -\mathbb{E}[b] \frac{C'_{XX}(t)}{\sigma_{X'}^2} \propto -C'_{XX}(t) \propto q(t), \quad (\text{E.4})$$

où un estimateur peut être construit par :

$$\hat{D}_{XX}^2(t) = \frac{1}{K} \sum_{k \geq 1} (X(t_k + t)|X(t_k) = a = 0, X(t_k) \geq 0), \quad (\text{E.5})$$

auquel cas nous savons, grâce à l'identité Eq. (1.6) que le décrément aléatoire estime la réponse impulsionnelle du système !

## E.2 Identification de l'oscillateur avec le décrément aléatoire

Présenté dans l'introduction, l'exemple de l'oscillateur amorti est repris en conclusion de ce manuscrit pour présenter un estimateur alternatif aux corrélateurs traditionnels. Rappelons



que l'oscillateur amorti est caractérisé par sa réponse impulsionnelle  $h(t)$  dont la fonction d'auto-corrélation est  $C_h(t)$ . La relation qui relie ces deux quantités correspond à l'identité de Ward pour un système à dissipation constante  $\eta$  :

$$q(t) = -\eta C'_h(t)u(t) \quad (\text{E.6})$$

où  $u(t)$  est l'échelon.

Nous allons confronter l'identification de  $C_q(t)$  et  $q(t)$  par le calcul de la corrélation de la mesure à l'utilisation de la méthode du décrément aléatoire.

La Fig. E.1 présente les signaux de simulation utilisés pour illustrer l'identité obtenue Eq. (1.6). La réponse impulsionnelle de l'oscillateur amorti est  $h(t)$  et ce dernier est excité par la source  $v(t)$ , variable aléatoire Gaussienne centrée et blanche. La réponse est  $q(t)$ . La mesure  $m(t)$  est la superposition de la réponse  $q(t)$  à une séquence de bruit d'acquisition i.i.d. Gaussienne centrée. Le RSB dans l'illustration vaut 2 dB.

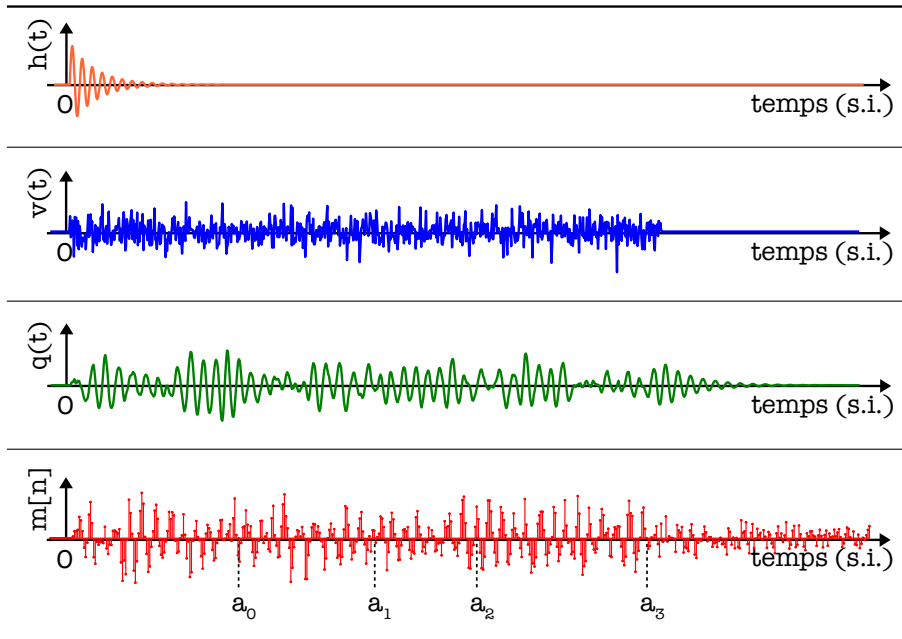


FIGURE E.1 – Identification d'un oscillateur amorti dans un contexte passif : signaux utilisés. Le filtrage de la source  $s(t)$  par la réponse impulsionnelle  $h(t)$  du circuit conduit à la réponse  $q(t)$ . Sa mesure est  $m[n]$ .

Nous confrontons l'estimation de  $C_h(t)$  et de  $h(t)$  par les méthodes de corrélation (corr) et de calcul de décrément aléatoire (rdd). Afin d'illustrer la notion de convergence, les quantités d'intérêt sont estimées à partir d'un nombre croissant de moyennages, utilisés pour approcher asymptotiquement l'opérateur d'espérance. L'estimation de  $C_q(t)$  est illustrée dans la Fig. E.2 et celle de  $q(t)$  dans la Fig. E.3. Les performances d'estimation des paramètres  $\eta$  et  $\omega$  sont présentées lorsque le RSB vaut 2 dB. Ces performances représentent le RMSE relatif (en %)

issu de la régression paramétrique des fonctions estimées, itéré 100 fois :

$$\text{rmser}(\hat{\bullet}) = \sqrt{\mathbb{E} [(\hat{\bullet} - \bullet)^2]} / \bullet \quad (\text{E.7})$$

Tout d'abord, il apparaît clair que les estimateurs proposés sont validés :

- L'identité Eq. (1.6) est vérifiée expérimentalement ;
- La méthode du décrément aléatoire permet bien d'estimer la corrélation  $C_q(t)$  et la réponse  $q(t)$ , sans utiliser l'identité Eq. (1.6).

La pulsation propre semble plus « facile » à estimer que l'atténuation. A première vue, l'estimateur par corrélation converge plus rapidement que celui par décrément aléatoire. Ceci se justifie en rappelant que ce dernier n'effectue aucune opération sur le signal lui même (il moyenne les réalisations extraites) tandis que la corrélation réalise un produit scalaire mettant en exergue, dès le premier extrait, les propriétés statistiques du signal. Cependant, il est clair que la complexité algorithmique n'est pas du même ordre de grandeur (produits contre sommes).

De plus, il apparaît que très peu de données sont requises pour permettre aux estimateurs d'atteindre un niveau de convergence satisfaisant pour estimer de manière robuste les paramètres du systèmes (pulsation, amortissement).

### E.3 Corrélateur hybride et peu complexe

Le décrément aléatoire est un corrélateur hybride dans le sens où le signal utilisé pour déterminer les origines de segments peut être réduit à son signe, c'est-à-dire à sa quantification sur 1 bit. Un exemple d'implémentation est proposé dans la Fig. E.4. Dans la conclusion de ce manuscrit, la proposition est faite d'observer le comportement du décrément aléatoire sur les signes de champs gaussiens corrélés. Ce faisant, il s'agirait de combiner la complexité très faible du calcul du décrément aléatoire, avec le contrôle sur la quantité de données manipulée par la quantification très forte des champs.

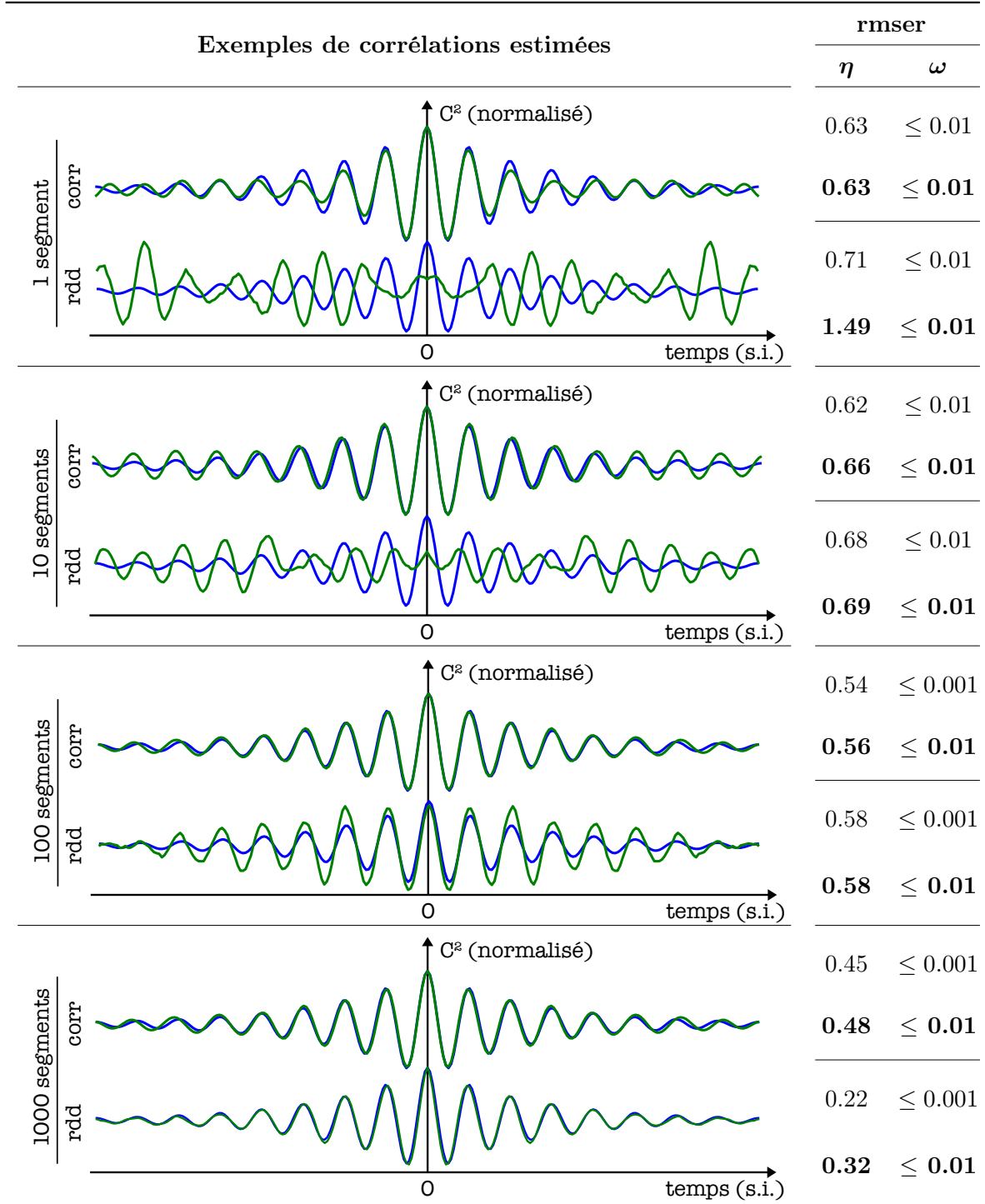


FIGURE E.2 – Estimation de la corrélation  $C_q(t)$  avec  $\{1, 10, 100, 1000\}$  segments extraits de la réponse  $m(t)$ . (Bleu)  $C_q(t)$  théorique. (Vert)  $\hat{C}_q(\tau)$ , estimateur par corrélation (corr) et estimateur par décrément aléatoire (rdd). Valeurs des rmser pour un RSB valant 2dB (normal) et -20dB (gras).

Exemples de réponses impulsionnelles estimées		rmser	
		$\eta$	$\omega$
<div>1 segment</div>	0.63	$\leq 0.01$	
	<b>0.64</b>	$\leq 0.01$	
	0.68	$\leq 0.01$	
	<b>2.34</b>	$\leq 0.01$	
<div>10 segments</div>	0.62	$\leq 0.01$	
	<b>0.65</b>	$\leq 0.01$	
	0.74	$\leq 0.01$	
	<b>1.07</b>	$\leq 0.01$	
<div>100 segments</div>	0.53	$\leq 0.001$	
	<b>0.60</b>	$\leq 0.01$	
	0.44	$\leq 0.001$	
	<b>0.57</b>	$\leq 0.01$	
<div>1000 segments</div>	0.43	$\leq 0.001$	
	<b>0.45</b>	$\leq 0.01$	
	0.14	$\leq 0.001$	
	<b>0.17</b>	$\leq 0.01$	

FIGURE E.3 – Estimation de la réponse impulsionnelle  $q(t)$  avec  $\{1, 10, 100, 1000\}$  segments extraits de la réponse  $m(t)$ . (Bleu)  $q(t)$  théorique. (Vert)  $\hat{q}(\tau)$ , estimateur par corrélation (corr) et estimateur par décrément aléatoire (rdd). Valeurs des rmser pour un RSB valant 2dB (normal) et -20dB (gras).

---

```

def random_decrement(x, y, N, M, *args, **kwargs):
    display = bool(kwargs.get('display', False))
    method = str(kwargs.get('method', False))
    threshold = float(kwargs.get('threshold', 0.01))

    # find trigger conditions indices
    if method == 'CIR': # CIR retrieval
        tmp = np.sign(y[:-N - 1] + threshold)
        tmp = np.diff(tmp)
    else: # IR retrieval
        tmp = np.sign(y[:-N - 1])
        tmp = np.diff(tmp > 0)

    zero_crossings = np.where(tmp)[0]
    zero_crossings = map(lambda a: a + 1, zero_crossings)
    zero_crossings = zero_crossings[:M]

    # extract N-len segments and average
    seg = np.zeros((len(zero_crossings), N))
    for i in range(len(zero_crossings)):
        start = int(zero_crossings[i])
        seg[i, :] = x[start:start + N]

    return np.mean(seg, axis=0) * (-1) ** np.int(method == 'CIR')

```

---

FIGURE E.4 – Fonction python de décréement aléatoire

---

**Résumé** — L'identité de Ward est une relation fonctionnelle qui permet d'identifier un milieu de propagation linéaire dissipatif. Dans les travaux exposés, cette identité est utilisée pour proposer de nouveaux modèles d'observation caractérisant un contexte d'estimation qualifié de passif : les sources qui excitent le système ne sont pas contrôlées par l'utilisateur. La théorie de l'estimation/détection dans ce contexte est étudiée et des analyses de performances sont menées sur divers estimateurs. La portée applicative des méthodes proposées concerne le domaine du Structural Health Monitoring (SHM), c'est-à-dire le suivi de l'état de santé des bâtiments comme un pont ou un immeuble. La modalité choisie est l'acoustique audible, cette dernière s'avérant complémentaire des techniques de l'état de l'art du SHM et permettant entre autre, d'accéder à des paramètres structuraux et géométriques. Des scénarios sont illustrés par la mise en oeuvre expérimentale des algorithmes développés et adaptés à des contraintes de calculs embarqués sur un réseau de capteurs autonome.

**Mots clés** : Estimation/détection à partir de sources d'opportunité — identification de systèmes linéaires dissipatifs — estimation bayésienne — estimation de paramètres structurels et géométriques.

---

---

**Abstract** — Ward identity is a functional relationship that enables damped linear system identification. This identity is used to provide new observation models that are characteristic of an estimation context where sources are uncontrolled by the user. An estimation and detection theory is derived from these models and various performances studies are conducted for several estimators. The reach of the proposed methods is extended to Structural Health Monitoring (SHM), that aims at measuring and tracking the health of buildings, such as a bridge or a sky-scraper for instance. The acoustic modality is chosen as it provides complementary parameters estimation to the state of the art in SHM, such as structural and geometrical parameters recovery. Some scenarios are experimentally illustrated by using the developed algorithms, adapted to fit the constraints set by embedded computation on an autonomous sensor network.

**Keywords** : Estimation/detection using sources of opportunity — damped linear system identification — bayesian estimation — structural and geometrical parameter estimation.

---